



Agitación para estimación anticipada de agitación pacientes de Alzheimer

Avril Simone Rojas Ríos

David Alejandro Gómez Torres

Santiago Torres Taborda

Universidad Ean

Ingeniería

Ingeniería de sistemas

Bogotá, Colombia

02/06/2026

Aplicación para estimación anticipada de agitación pacientes de Alzheimer

Avril Simone Rojas Ríos

David Alejandro Gómez Torres

Santiago Torres Taborda

Trabajo de grado presentado como requisito para optar al título de:

Ingeniería de sistemas.

Director (a):

Alexander Garcia

Universidad Ean

Ingeniería

Ingeniería de sistemas

Bogotá, Colombia

02/06/2026

Nota de aceptación:

Firma del jurado

Firma del jurado

Firma del director del trabajo de grado

Bogotá, 04/06/20

Resumen

Este proyecto busca desarrollar una aplicación tecnológica basada en el análisis de datos para predecir episodios de agitación en pacientes con Alzheimer, transformando el cuidado reactivo en preventivo. El problema central es la agitación disruptiva que agota a los cuidadores y afecta la calidad de vida del paciente. Los objetivos incluyen definir variables conductuales en un registro digital estandarizado, implementar un modelo de aprendizaje automático para estimar riesgos y evaluar su desempeño mediante pruebas piloto. Se analiza la viabilidad técnica considerando la privacidad y sostenibilidad. La investigación parte de la hipótesis de que un registro digital organizado reduce la carga mental del cuidador y mejora su seguridad al manejar la enfermedad. Integrando la ciencia de datos y la salud en el contexto colombiano, el estudio busca anticipar crisis para facilitar intervenciones tempranas que mejoren la gestión del cuidado y el bienestar de las familias afectadas.

Palabras clave: Alzheimer, Agitación, Modelo

Abstract

This project proposes to develop a technological application based on data analysis to predict agitation episodes in Alzheimer's patients, changing from reactive to preventive care. The root cause is disruptive agitation that exhausts caregivers and affects the patient's quality of life. Objectives include defining behavioral variables in a standardized digital record, implementing a machine learning model for risk estimation, and evaluating its performance through pilot tests. Technical viability is analyzed considering privacy and sustainability. The research is based on the hypothesis that an organized digital record reduces the caregiver's mental burden and enhances their sense of security in managing the disease. By integrating data science and healthcare within the Colombian context, the study seeks to anticipate crises to facilitate early interventions that improve care management and the well-being of affected families.

Keywords: Alzheimer, Agitation, Model

Introducción

La enfermedad de Alzheimer es actualmente uno de los mayores retos de salud pública y social del mundo. Este problema no solo se debe a la pérdida de memoria de los pacientes, sino principalmente a los cambios de conducta que ocurren durante la enfermedad (World Health Organization, 2023). Entre estos cambios, la agitación es uno de los más difíciles de manejar, ya que suele manifestarse como gritos o conductas físicas agresivas que afectan la tranquilidad del hogar. Por mucho tiempo, la forma de enfrentar estos momentos ha sido reactiva; es decir, usualmente solo se actúa cuando el paciente ya está alterado (Homdee et al., 2019). En consecuencia, es necesario cambiar hacia un modelo preventivo donde se use la tecnología para detectar estas crisis antes de que sucedan.

En el mundo científico ya se está estudiando cómo el uso de dispositivos inteligentes y el análisis de datos pueden ayudar; por ejemplo, diferentes investigadores han propuesto que es posible identificar señales en el cuerpo o en el movimiento que avisan cuando alguien está por alterarse (Paul Ekman, 2003). Diversos estudios han demostrado que variables como la calidad del sueño y los cambios en el ritmo cardíaco pueden ser utilizadas por modelos de aprendizaje automático para anticipar alteraciones conductuales (Ramesh et al., 2025). Sin embargo, todavía existe un vacío importante: no hay herramientas sencillas que los cuidadores puedan usar con facilidad en sus casas. Se ha identificado que, para que estas herramientas sean útiles en el hogar, deben ser interpretables y basarse en actividades de la vida diaria, capturadas de forma no invasiva (Bafaloukou et al., 2024; Fletcher-Lloyd et al., 2021).

Esta investigación se encuentra en un punto en el que se unen la ciencia de datos, la tecnología en salud y el cuidado de los adultos mayores. El problema principal es que la agitación suele verse como algo que ocurre de repente, pero en realidad es el resultado de pequeñas señales que no se percibieron a tiempo. Actualmente, no existe una forma organizada y digital de anotar estas señales, por esta razón, el estudio se enfoca en crear una solución tecnológica que junte toda esta información en un modelo capaz de predecir con tiempo a los cuidadores para que puedan intervenir antes de que la situación se salga de control. (HekmatiAthar et al., 2021)

De acuerdo con lo anterior, el propósito del proyecto es desarrollar una aplicación tecnológica que utilice técnicas de análisis de datos para estimar el riesgo de agitación en pacientes con Alzheimer y apoyar a los cuidadores en la toma de decisiones, lo que hará el cuidado más llevadero. Para lograrlo, se definirán qué señales de conducta son las más importantes para anotarlas de forma digital, después se creará un modelo de aprendizaje automático que analice esos datos y diga qué tan probable es que ocurra una crisis y finalmente, se harán pruebas para ver qué tan bien funciona la aplicación y si es fácil de usar para las personas, revisando además si es una herramienta segura y que se pueda mantener funcionando a largo plazo.(Zhang et al., 2022)

A partir de este panorama, surge la pregunta central de la investigación: ¿Es posible crear una herramienta tecnológica que analice conductas y utilice el aprendizaje automático para detectar con precisión y anticipación estados de crisis y agitación, ayudando a los cuidadores a actuar a tiempo? Esta pregunta busca entender si el uso de datos puede realmente ayudar a las personas más allá de lo que ven con sus propios ojos, mejorando así la calidad de vida del paciente y su entorno familiar (Oliveira et al., 2022)

Como respuesta a esta pregunta, se plantea la hipótesis de que implementar un modelo de registro digital organizado disminuye significativamente el cansancio mental de quien cuida y aumenta su sensación de seguridad al manejar la enfermedad. Se cree que esto pasará sin importar qué tan compleja sea la tecnología interna de la aplicación, siempre y cuando la forma de anotar los datos sea clara y sencilla. Esta idea se pondrá a prueba durante el trabajo para confirmar si la tecnología propuesta es la mejor solución para cuidar a los pacientes de forma preventiva y proactiva.

Marco teórico

La demencia es una enfermedad neurodegenerativa caracterizada por la pérdida progresiva de capacidades cognitivas como la memoria, el pensamiento y la comprensión, originada por la muerte de neuronas en áreas críticas como el hipocampo y el córtex cerebral (Organización Mundial de la Salud [OMS], 2025), una de las funciones que se ven afectadas por el daño de estas áreas es el lenguaje, que, en conjunto con un deterioro físico, generan al paciente una dependencia total. En este contexto, uno de los síntomas más relevantes en el desarrollo y comportamiento de los pacientes es la agitación, esta es definida como el exceso de actividades vocales y motoras que representan un estado de irritación o disgusto (Meunier et al., 2024). Estas crisis de comportamiento se deben a las molestias físicas y emocionales que los pacientes son incapaces de comunicar debido a su deterioro físico y cognitivo (HekmatiAthar et al., 2021).

Para comprender por qué surgen estas conductas se utiliza el Modelo de Comportamiento Comprometido por la Demencia impulsado por Necesidades o más conocido como NDB (Need-driven Dementia-compromised Behavior), que explica cómo la agitación ocurre cuando factores proximales (entorno, cuidadores) y factores de fondo (salud, historial clínico) interactúan entre sí (Algase et al., 1996), por ejemplo, el síndrome vespertino es un estado en el que la agitación se ve aumentada al anochecer por la fatiga acumulada durante el día (Fletcher-Lloyd et al., 2021). Por lo anterior, uno de los puntos más importantes para la investigación y práctica clínica es la clasificación de estos comportamientos para poder cuantificarlos, las categorías que componen esta clasificación son la agresión física, conductas físicas no agresivas y agitación verbal, este sistema se conoce como Inventario de Agitación de Cohen-Mansfield (CMAI) (Meunier et al., 2024).

Desde una perspectiva colombiana, el Alzheimer afecta a entre 250,000 y 340,000 ciudadanos, con un promedio de edad mayor o igual a 60 años (Ministerio de Salud y Protección Social, 2020). Además, en el país se cuenta con una mutación genética única en el mundo, conocida como PSEN1 (G280A) o mutación paisa, que se presenta en familias del municipio de Yarumal, Antioquia, y caracteriza por su inicio temprano, entre los 35 y 40 años. La mutación paisa fue descubierta y nombrada por el doctor Francisco Lopera Restrepo, quien fue una figura muy importante en el desarrollo de investigaciones y proyectos para conocer a mayor profundidad la enfermedad, su trabajo fue lo suficientemente relevante para atraer la atención de la comunidad internacional y que estos contribuyeran a dichas investigaciones (Gómez, 2024).

Si se habla del impacto de la tecnología, los proyectos más importantes se enfocan en el monitoreo de la enfermedad, la ralentización del deterioro y el diagnóstico a una edad temprana, estas herramientas tienen como principales recursos el uso de inteligencia artificial (IA), machine learning y digital twins (Bafaloukou et al., 2024; Bertolini et al., 2020; Ramesh et al., 2025). Unos de los proyectos más conocidos es AngelSense, este es un dispositivo GPS para el rastreo de niño y adultos con necesidades especiales, el GPS se sujeta a la ropa del usuario con sujetadores que solo el cuidador puede quitar, ofreciéndole a este mismo una localización continua, por ejemplo, el cuidador puede delimitar zonas no seguras o desconocidas que generaran una alarma cuando el paciente se encuentre en ellas, además de incluir un altavoz por el cual se puede comunicar con el usuario y escuchar su entorno (Konexión Alzheimer, 2026), muchos de sus resultados son positivos resaltando la tranquilidad que genera en los familiares y cuidadores, sin embargo algunas de sus críticas se dirigen al rendimiento tecnológico del dispositivo, siendo puntos bajos su batería o componentes de una calidad básica que pueden presentar fallas con facilidad. Por último, otro proyecto de gran relevancia es CogniFit, una plataforma con entrenamientos y evaluaciones cognitivas, entre los que se incluyen juegos mentales personalizados que miden, entrenan y mejoran habilidades como la memoria, la atención y la concentración (Konexión Alzheimer, 2026). Este proceso inicia con una prueba que busca medir las fortalezas y debilidades de los pacientes. Con esto, se crean rutinas de juegos mentales que se ajustan al rendimiento del usuario. Los resultados de estos juegos permiten monitorear a los pacientes por medio de informes y métricas de evolución.

Por último, uno de los puntos más importantes es el uso de un modelo de estimación de crisis de agitación en pacientes con Alzheimer; aquí se buscan y analizan proyectos de inteligencia artificial y machine learning para la salud, en el área de anticipación de deterioro cognitivo o conductas disruptivas. Los modelos que cumplen con las características y funciones necesarias son aquellos que usen como estadísticas: variables clínicas, conductuales o sensoriales, que tengan como clasificación de resultados: binaria (riesgo alto/bajo), multiclase (alto/medio/bajo) o incluso regresión lineal (puntuación de riesgo continua), y que apliquen ML para la predicción de demencia, síntomas neuropsiquiátricos o agitación, con datos clínicos o digitales. Con las características y funciones de los modelos ya definidas, existen dos tipos de modelos que cumplen dichos requerimientos. Uno de estos es Random Forest, más conocido por ser un modelo de clasificación supervisada que se usa para predecir el riesgo de demencia, al manejar un conjunto de árboles de decisión con variables mixtas y que evita sobreajustes, para votar por un resultado. Otro modelo importante son las redes neuronales artificiales; se conforman por una capa que aprende patrones de comportamiento de variables como el sueño o la actividad física de un paciente; por lo tanto, se utilizan en trabajos de predicción de síntomas neuropsiquiátricos con datos digitales y hacen parte de los modelos de Deep learning. Para determinar la efectividad del modelo, se usarán métricas básicas de clasificación, como la precisión (accuracy) que es la proporción de predicciones correctas sobre el total, la sensibilidad (recall) que es el porcentaje de episodios de agitación que fueron correctamente identificados, métricas clínicas como el valor predictivo positivo (VPP) que es la proporción de alertas de riesgo alto de crisis que realmente sucedieron y métricas de discriminación como el AUC-ROC (área bajo la curva ROC) que mide la capacidad del modelo para discriminar entre un riesgo alto y bajo.

Metodología

Este proyecto tiene un enfoque cuantitativo, con diseño no experimental y alcance descriptivo-correlacional, ya que no se manipularán variables, se observará el fenómeno de forma natural y se realizarán mediciones a través del tiempo que nos permitan evaluar el correcto funcionamiento del

aplicativo y sus registros conductuales (Fletcher-Lloyd et al., 2021), para su posterior análisis y comprender mejor los datos asociados a crisis de comportamiento en pacientes de Alzheimer.

La población de estudio estará compuesta por pacientes diagnosticados con enfermedad de Alzheimer en Colombia, cuyos cuidadores participarán en el registro sistemático de variables conductuales mediante un instrumento digital estructurado. Para el entrenamiento del modelo se dispondrán datos simulados generados por inteligencia artificial y datos de bases de datos públicas. Los datos que se van a utilizar tienen que evitar valores atípicos que puedan distorsionar los resultados, tienen que provenir de fuentes fiables o reconocidas, tener un mínimo de valores faltantes y presentar lógica entre variables, estos criterios permiten que los datos que se van a utilizar sean confiables y representativos.

Además de establecer criterios de legibilidad para los datos que se utilizarán, se considera incorporar datos e información de otros países; estos se obtendrán de repositorios públicos y estudios internacionales acerca de la demencia senil y el Alzheimer. Al incluir estos datos, se amplía la diversidad de registros utilizados en el entrenamiento del modelo, además, al poder comparar contextos, los datos internacionales se estandarizarán y adaptarán, especialmente los datos de frecuencia de conductas agitadas y variables ambientales.

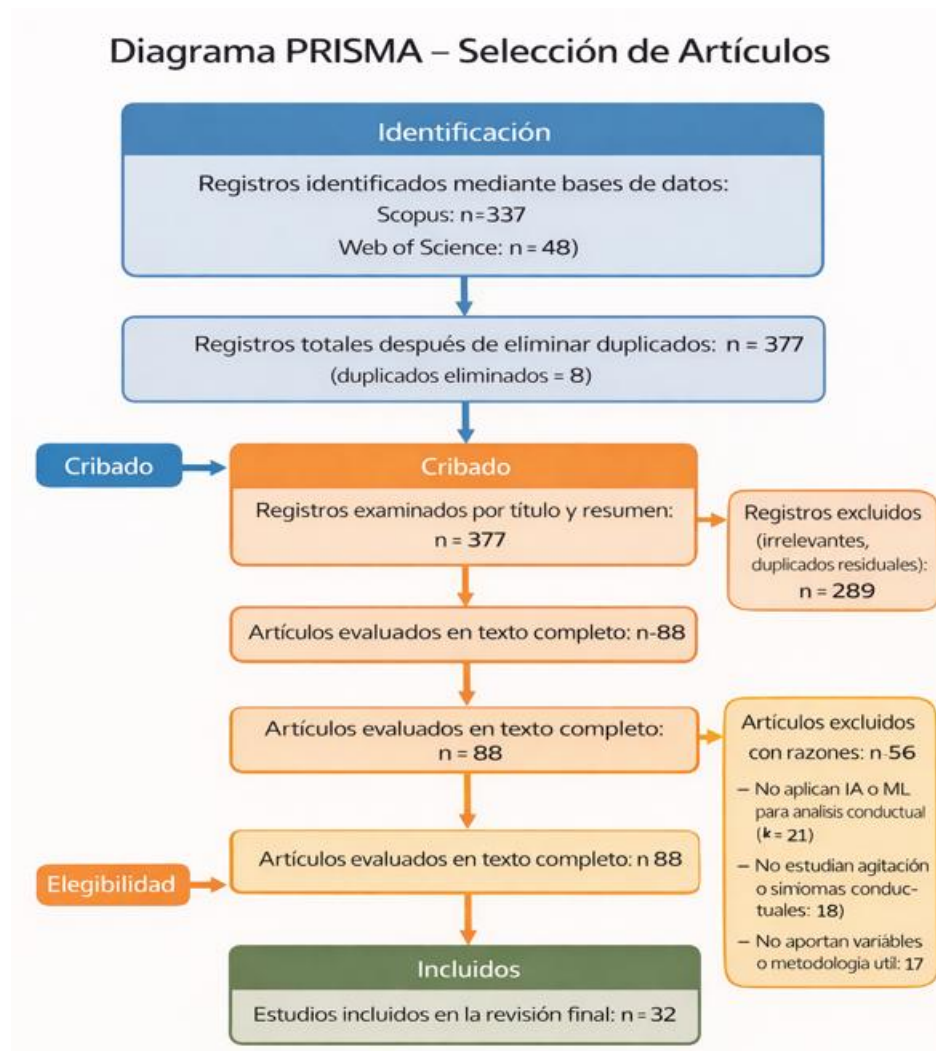
Como instrumento de medición se empleará una adaptación de la escala Cohen-Mansfield Agitation Inventory (CMAI), un instrumento validado internacionalmente para cuantificar la frecuencia de conductas disruptivas (Meunier et al., 2024), complementada con variables ambientales y conductuales (como la inquietud motora o la agitación verbal).

Para el análisis de la información se aplicarán técnicas de estadística descriptiva y análisis correlacional, con el fin de identificar patrones asociados a la aparición de episodios de agitación. Posteriormente, se desarrollará un modelo predictivo basado en técnicas de aprendizaje supervisado para estimar el riesgo de ocurrencia de dichos episodios. El análisis se realizará mediante herramientas computacionales especializadas.

Estado del arte

Figura 1

Diagrama PRISMA para selección de artículos



Nota: Información obtenida por los autores a partir del desarrollo de la metodología PRISMA

La búsqueda bibliográfica se llevó a cabo en las bases de datos Scopus y Web of Science, utilizando inicialmente fórmulas de búsqueda amplias relacionadas con la enfermedad de Alzheimer, agitación, demencia, síntomas conductuales, técnicas de inteligencia artificial, machine learning y modelos predictivos.

En Scopus se utilizó la siguiente ecuación de búsqueda: “(Alzheimer OR dementia) AND (agitation OR "behavioral symptoms" OR "neuropsychiatric symptoms") AND ("machine

learning" OR "artificial intelligence" OR prediction OR monitoring)” a diferencia de Web of Science en la que por flexibilidad de sintaxis utilizamos “TITLE-ABS-KEY(("Alzheimer's disease" OR dementia) AND (agitation OR "behavioral symptoms" OR "neuropsychiatric symptoms") AND ("machine learning" OR "artificial intelligence" OR prediction OR monitoring OR sensor OR wearable OR "activity recognition"))”.

Posteriormente, se aplicó un proceso de filtrado basado en criterios de año de publicación (2020-2026), tipo de documentación (“article” y “review article”), idioma (inglés y español) y por último áreas temáticas relevantes (“medicine”, “neuroscience”, “psychology”, “computer science”, “pharmacology, Toxicology and Pharmaceutics”, “engineering” y “environmental science”) respectivamente.

Tras el proceso de filtrado, se obtuvieron 337 registros en Scopus y 48 registros en Web of Science, para un total de 385 registros identificados en la fase de búsqueda detallada. Consecuentemente, se realizó la eliminación de duplicados al tratarse de dos bases de datos distintas. Se encontraron 8 registros duplicados, dando como resultado 377 registros únicos.

En la fase de cribado, se revisaron los títulos y resúmenes de estos 377 registros, excluyendo aquellos estudios que no se relacionaran directamente con el objetivo del proyecto. Trabajos centrados exclusivamente en farmacología, genética, biomarcadores no conductuales, diagnóstico general de demencia sin aplicación tecnológica, o investigaciones que no utilicen inteligencia artificial, sensores, monitoreo o modelos predictivos aplicados a síntomas conductuales.

Para la fase de elegibilidad, se evaluaron de forma detallada los documentos que superaron el cribado inicial. Se excluyeron aquellos estudios que, aunque están relacionados con demencia o Alzheimer, no aportan variables, metodología, arquitectura tecnológica o resultados útiles para el desarrollo de una solución orientada a la estimación anticipada de episodios de agitación.

Finalmente se seleccionaron 32 estudios para la síntesis cualitativa que son la base del estado del arte.

Las investigaciones más recientes acerca del Alzheimer y la demencia senil han sufrido una transición significativa, pasando de un enfoque centrado en diagnósticos clínicos y tratamientos farmacológicos hacia el uso de tecnologías disruptivas como inteligencia artificial, sensores, monitoreo remoto y análisis de comportamiento (Addae et al., 2024; Fernandes et al., 2025). Este cambio ha permitido no solo mejorar los procesos médicos, sino también avanzar hacia modelos de cuidado más preventivos y personalizados. En este contexto, la agitación deja de ser vista únicamente como una consecuencia del deterioro cognitivo y comienza a considerarse un fenómeno medible, caracterizable digital y potencialmente predecible (Malioukis et al., 2026; Kleine Deters et al., 2024), lo cual resulta clave para el desarrollo de herramientas orientadas al cuidado cotidiano de los pacientes.

Un primer grupo de investigaciones se centra en el uso de inteligencia artificial aplicada a la demencia. La mayoría de los estudios validan su utilidad en cuatro aspectos principales: detección temprana de la enfermedad, monitoreo de síntomas conductuales, personalización del cuidado y apoyo a cuidadores (Ranson et al., 2023; Xie et al., 2020). Asimismo, se evidencia que la incorporación de variables conductuales y ambientales mejora la capacidad predictiva de los modelos, permitiendo anticipar cambios en la evolución del paciente (Mar et al., 2020; Gill et al., 2020). Sin embargo, estos estudios también destacan limitaciones importantes como el tamaño reducido de las muestras, dificultades en la validación clínica y la baja transferencia de modelos a entornos reales (Addae et al., 2024).

El segundo grupo de investigaciones se enfoca en la predicción directa de crisis de agitación o síntomas relacionados. Diversos demuestran que la agitación no ocurre de forma aleatoria, sino que responde a patrones identificables a partir de señales fisiológicas, conductuales y ambientales (Ramesh et al., 2025; Fletcher-Lloyd et al., 2021). Variables como la calidad del sueño, la actividad diaria, los cambios fisiológicos y el comportamiento motor han mostrado ser indicadores relevantes para anticipar episodios de agitación (Hasan et al., 2025; Rykov et al., 2024). Estos

hallazgos sustentan científicamente la viabilidad de modelos predictivos basados en aprendizaje automático, los cuales permiten estimar riesgos antes de que ocurran crisis.

Por último, un tercer grupo de estudios se enfoca en el uso de dispositivos tecnológicos para el monitoreo continuo de pacientes. En este ámbito, se destacan herramientas como sensores, biomarcadores digitales y dispositivos wearables, los cuales permiten capturar información en tiempo real sobre el comportamiento del paciente (Iaboni et al., 2022; Liu et al., 2024). Estas tecnologías han demostrado ser capaces de registrar variables como movimiento, patrones de actividad, ubicación y cambios en la rutina diaria (Arunachalam et al., 2023; Kim et al., 2022), lo que permite obtener datos objetivos y continuos, reduciendo la dependencia de observaciones subjetivas por parte de los cuidadores (Rezvani et al., 2021; Favela et al., 2020).

En conjunto, la literatura revisada permite identificar tres tendencias principales. En primer lugar, existe una creciente aceptación del uso de modelos sistemáticos, machine learning e inteligencia artificial para el análisis de la demencia y sus síntomas (Addae et al., 2024). En segundo lugar, los sensores y dispositivos wearables se consolidan como fuentes clave de datos conductuales (Cheung et al., 2022; Malioukis et al., 2026). En tercer lugar, la agitación ha comenzado a abordarse desde una perspectiva predictiva, aunque muchos estudios aún se desarrollan en entornos controlados o institucionales (Bafaloukou et al., 2025; Kleine Deters et al., 2024).

A pesar de estos avances, persisten vacíos importantes que justifican el desarrollo de nuevas soluciones tecnológicas. Muchos estudios se enfocan en detectar la agitación cuando ya está ocurriendo, pero son menos los que abordan la estimación anticipada en contextos domésticos reales (Bafaloukou et al., 2025). Además, varias soluciones dependen de sensores especializados o infraestructuras complejas que dificultan su implementación en el hogar (Arunachalam et al., 2023).

Por lo tanto, el análisis del estado del arte respalda la pertinencia del proyecto al evidenciar que los datos conductuales, fisiológicos y de actividad diaria pueden utilizarse para anticipar síntomas

relevantes en demencia. Sin embargo, aún existe una brecha entre la investigación experimental y el desarrollo de herramientas prácticas orientadas a cuidadores. En este contexto, el presente proyecto busca contribuir al desarrollo de un sistema capaz de estimar de forma anticipada episodios de agitación en personas con Alzheimer, integrando técnicas de inteligencia artificial y análisis de comportamiento para apoyar la toma de decisiones en el cuidado cotidiano.

Tabla 1

Resultados obtenidos en Scopus mediante ecuación de búsqueda

Autor(es) / Año	Título del estudio	Objetivo principal	Metodología empleada	Variables analizadas	Principales hallazgos	Limitaciones reportadas	Relación con el proyecto actual
Fletcher-Lloyd et al. (2021)	Home monitoring of daily living activities and prediction of agitation risk in a cohort of people living with dementia	Predecir riesgo, presencia o severidad de demencia, agitación o síntomas neuropsiquiátricos a partir de datos clínicos, conductuales o sensoriales.	Estudio aplicado en salud digital y demencia con análisis computacional.	Agitación	El estudio reporta viabilidad para predecir riesgo o severidad de síntomas/deterioro con modelos de IA a partir de variables conductuales, clínicas o sensoriales.	Tamaño de muestra o generalización limitada.	Muestra cómo captar señales de agitación con sensores o monitoreo continuo; útil para variables e instrumentación.
Ramesh et al. (2025)	Machine Learning Prediction of Agitation in Dementia Patients Using Sleep and Physiological Data	Predecir riesgo, presencia o severidad de demencia, agitación o síntomas neuropsiquiátricos a partir de datos clínicos, conductuales o sensoriales.	Estudio aplicado con modelos de machine learning, deep learning o IA sobre datos clínicos, conductuales o digitales.	Sueño, señales fisiológicas y agitación	El estudio reporta viabilidad para predecir riesgo o severidad de síntomas/deterioro con modelos de IA a partir de variables conductuales, clínicas o sensoriales.	Requerimiento de datos suficientes y bien etiquetados.	Aporta un antecedente directo de predicción/detección de agitación con IA, muy cercano al objetivo del proyecto.
Rykov et al. (2024)	Predicting the severity of mood and neuropsychiatric symptoms from digital biomarkers using wearable physiological data and deep learning	Predecir riesgo, presencia o severidad de demencia, agitación o síntomas neuropsiquiátricos a partir de datos clínicos, conductuales o sensoriales.	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales; uso de IA o ML para el análisis.	Señales fisiológicas, síntomas neuropsiquiátricos, datos de wearables / acelerómetros, biomarcadores digitales	El estudio reporta viabilidad para predecir riesgo o severidad de síntomas/deterioro con modelos de IA a partir de variables conductuales, clínicas o sensoriales.	Dependencia de calidad/calibración de sensores; requerimiento de datos suficientes y bien etiquetados.	Aporta un antecedente metodológico o tecnológico relacionado con monitoreo, predicción o análisis conductual en demencia.
Bafaloukou et al. (2025)	An interpretable machine learning tool for in-home monitoring of agitation episodes in people living with dementia: a proof-of-concept study	Monitorear síntomas conductuales o cambios en la rutina diaria de personas con demencia en contextos reales.	Estudio de prueba de concepto con validación inicial.	Agitación	El monitoreo continuo permite capturar cambios de rutina y síntomas conductuales de interés clínico en contextos reales.	Validación preliminar y muestra limitada; requerimiento de datos suficientes y bien etiquetados.	Aporta un antecedente directo de predicción/detección de agitación con IA, muy cercano al objetivo del proyecto.

Kleine Deters et al. (2024)	Sensor-based agitation prediction in institutionalized people with dementia A systematic review	Sintetizar la evidencia disponible sobre tecnologías, sensores o IA aplicadas a demencia, síntomas conductuales o agitación.	Revisión sistemática de literatura.	agitación, datos de sensores / IoT	La literatura revisada muestra potencial de la IA, los sensores y el monitoreo digital para detectar, seguir o anticipar síntomas en demencia, aunque con alta heterogeneidad metodológica.	Heterogeneidad de estudios y métricas; dependencia de calidad/calibración de sensores.	Sintetiza evidencia sobre predicción de agitación con sensores; sirve para justificar el problema y la brecha tecnológica.
Iaboni et al. (2022)	Wearable multimodal sensors for the detection of behavioral and psychological symptoms of dementia using personalized machine learning models	Detectar o clasificar comportamientos, síntomas o episodios relevantes en personas con demencia mediante análisis computacional.	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales; uso de IA o ML para el análisis.	Síntomas conductuales / BPSD, datos de wearables / acelerometría, datos de sensores / IoT	Se demuestra factibilidad para detectar o clasificar episodios y comportamientos relevantes mediante IA y datos digitales.	Dependencia de calidad/calibración de sensores; requerimiento de datos suficientes y bien etiquetados.	Evidencia que los síntomas conductuales pueden detectarse con sensores/wearables y modelos personalizados.
Baidjnath Misier et al. (2024)	Monitoring challenging behaviors in dementia with sensor technology	Monitorear síntomas conductuales o cambios en la rutina diaria de personas con demencia en contextos reales.	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales; uso de IA o ML para el análisis.	Datos de sensores / IoT	El monitoreo continuo permite capturar cambios de rutina y síntomas conductuales de interés clínico en contextos reales.	Dependencia de calidad/calibración de sensores.	Aporta un antecedente metodológico o tecnológico relacionado con monitoreo, predicción o análisis conductual en demencia.
Arunachalam et al. (2023)	A smart Alzheimer's patient monitoring system with IoT-assisted technology through enhanced deep learning approach	Monitorear síntomas conductuales o cambios en la rutina diaria de personas con demencia en contextos reales.	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales; uso de IA o ML para el análisis.	Datos de sensores / IoT	El monitoreo continuo permite capturar cambios de rutina y síntomas conductuales de interés clínico en contextos reales.	Dependencia de calidad/calibración de sensores; requerimiento de datos suficientes y bien etiquetados.	Aporta arquitectura de monitoreo remoto e integración tecnológica útil para el diseño de la aplicación.
Kim et al. (2022)	IoT-Based Unobtrusive Physical Activity Monitoring System for Predicting Dementia	Predecir riesgo, presencia o severidad de demencia, agitación o síntomas neuropsiquiátricos a partir de datos clínicos, conductuales o sensoriales.	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales; uso de IA o ML para el análisis.	Datos de sensores / IoT, actividad diaria / marcha	El estudio reporta viabilidad para predecir riesgo o severidad de síntomas/deterioro con modelos de IA a partir de variables conductuales, clínicas o sensoriales.	Dependencia de calidad/calibración de sensores.	Aporta arquitectura de monitoreo remoto e integración tecnológica útil para el diseño de la aplicación.
Mar et al. (2020)	Validation of Random Forest Machine Learning Models to Predict Dementia-Related Neuropsychiatric Symptoms in Real-World Data	Predecir riesgo, presencia o severidad de demencia, agitación o síntomas neuropsiquiátricos a partir de datos clínicos, conductuales o sensoriales.	Estudio aplicado con modelos de machine learning, deep learning o IA sobre datos clínicos, conductuales o digitales.	Síntomas neuropsiquiátricos	El estudio reporta viabilidad para predecir riesgo o severidad de síntomas/deterioro con modelos de IA a partir de variables conductuales, clínicas o sensoriales.	Requerimiento de datos suficientes y bien etiquetados.	Aporta antecedentes de uso de IA/ML para detección o predicción en demencia.
Gill et al. (2020)	Using Machine Learning to Predict Dementia from Neuropsychiatric Symptom and Neuroimaging Data	Predecir riesgo, presencia o severidad de demencia, agitación o síntomas neuropsiquiátricos a partir de datos	Estudio aplicado con modelos de machine learning, deep learning o IA sobre datos	Neuroimagen, síntomas neuropsiquiátricos	El estudio reporta viabilidad para predecir riesgo o severidad de síntomas/deterioro con modelos de IA a partir	Requerimiento de datos suficientes y bien etiquetados.	Aporta antecedentes de uso de IA/ML para detección o predicción en demencia.

		clínicos, conductuales o sensoriales.	clínicos, conductuales o digitales.		de variables conductuales, clínicas o sensoriales.		
Malioukis et al. (2026)	Digital biomarkers for early agitation detection in dementia: a scoping review of emerging wearable and smart technologies for personalized care	Sintetizar la evidencia disponible sobre tecnologías, sensores o IA aplicadas a demencia, síntomas conductuales o agitación.	Revisión de alcance (scoping review).	Agitación, datos de wearables / acelerometría, biomarcadores digitales	La literatura revisada muestra potencial de la IA, los sensores y el monitoreo digital para detectar, seguir o anticipar síntomas en demencia, aunque con alta heterogeneidad metodológica.	Heterogeneidad de estudios y métricas; dependencia de calidad/calibración de sensores.	Muestra cómo captar señales de agitación con sensores o monitoreo continuo; útil para variables e instrumentación.
Jeon et al. (2023)	Early Alzheimer's Disease Diagnosis Using Wearable Sensors and Multilevel Gait Assessment: A Machine Learning Ensemble Approach	Explorar el uso de tecnologías digitales e inteligencia artificial para apoyar el seguimiento de demencia y síntomas asociados.	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales; uso de IA o ML para el análisis.	Datos de wearables / acelerometría, datos de sensores / IoT, actividad diaria / marcha	El trabajo aporta evidencia de que las tecnologías digitales pueden apoyar el seguimiento y análisis de síntomas en demencia.	Dependencia de calidad/calibración de sensores; requerimiento de datos suficientes y bien etiquetados.	Aporta un antecedente metodológico o tecnológico relacionado con monitoreo, predicción o análisis conductual en demencia.
Cheung et al. (2022)	Wrist accelerometry for monitoring dementia agitation behaviour in clinical settings: A scoping review	Sintetizar la evidencia disponible sobre tecnologías, sensores o IA aplicadas a demencia, síntomas conductuales o agitación.	Revisión de alcance (scoping review).	Agitación, datos de wearables / acelerometría	La literatura revisada muestra potencial de la IA, los sensores y el monitoreo digital para detectar, seguir o anticipar síntomas en demencia, aunque con alta heterogeneidad metodológica.	Heterogeneidad de estudios y métricas; dependencia de calidad/calibración de sensores.	Muestra cómo captar señales de agitación con sensores o monitoreo continuo; útil para variables e instrumentación.
Hasan et al. (2025)	Machine Learning to Detect Motor Agitation in People With Dementia Using Real-Time Location Data	Detectar o clasificar comportamientos, síntomas o episodios relevantes en personas con demencia mediante análisis computacional.	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales; uso de IA o ML para el análisis.	Localización en tiempo real, agitación	Se demuestra factibilidad para detectar o clasificar episodios y comportamientos relevantes mediante IA y datos digitales.	Requerimiento de datos suficientes y bien etiquetados.	Aporta un antecedente directo de predicción/detección de agitación con IA, muy cercano al objetivo del proyecto.
Addae et al. (2024)	Smart Solutions for Detecting, Predicting, Monitoring, and Managing Dementia in the Elderly: A Survey	Sintetizar la evidencia disponible sobre tecnologías, sensores o IA aplicadas a demencia, síntomas conductuales o agitación.	Artículo de revisión tipo survey.	Variables clínicas, variables conductuales, señales digitales	La literatura revisada muestra potencial de la IA, los sensores y el monitoreo digital para detectar, seguir o anticipar síntomas en demencia, aunque con alta heterogeneidad metodológica.	Heterogeneidad de estudios y métricas.	Aporta contexto general, tendencias, métodos y brechas del campo.
Yuan et al. (2024)	Modeling and mobile home monitoring of behavioral and psychological symptoms of dementia (BPSD)	Monitorear síntomas conductuales o cambios en la rutina diaria de personas con demencia en contextos reales.	Estudio aplicado en salud digital y demencia con análisis computacional.	Síntomas conductuales / BPSD, registros móviles	El monitoreo continuo permite capturar cambios de rutina y síntomas conductuales de interés clínico en contextos reales.	Tamaño de muestra o generalización limitada.	Aporta metodologías para monitoreo de síntomas conductuales en entorno real y apoyo al cuidador.
Ranson et al. (2023)	Harnessing the potential of machine learning and artificial intelligence for dementia research	Explorar el uso de tecnologías digitales e inteligencia artificial para apoyar el seguimiento de	Estudio aplicado con modelos de machine learning, deep learning o IA sobre datos	Variables clínicas, variables conductuales, señales digitales	El trabajo aporta evidencia de que las tecnologías digitales pueden apoyar el seguimiento y análisis	Requerimiento de datos suficientes y bien etiquetados.	Aporta antecedentes de uso de IA/ML para detección o predicción en demencia.

		demencia y síntomas asociados.	clínicos, conductuales o digitales.		de síntomas en demencia.		
Rezvani et al. (2021)	Analysing behavioural changes in people with dementia using in-home monitoring technologies	Monitorear síntomas conductuales o cambios en la rutina diaria de personas con demencia en contextos reales.	Estudio aplicado en salud digital y demencia con análisis computacional.	Síntomas conductuales / BPSD	El monitoreo continuo permite capturar cambios de rutina y síntomas conductuales de interés clínico en contextos reales.	Tamaño de muestra o generalización limitada.	Aporta un antecedente metodológico o tecnológico relacionado con monitoreo, predicción o análisis conductual en demencia.
Rangseekajee et al. (2021)	Mobile application for monitoring behavioral and psychological symptoms of dementia in patients with moderate to severe dementia	Monitorear síntomas conductuales o cambios en la rutina diaria de personas con demencia en contextos reales.	Estudio aplicado en salud digital y demencia con análisis computacional.	Síntomas conductuales / BPSD, registros móviles	El monitoreo continuo permite capturar cambios de rutina y síntomas conductuales de interés clínico en contextos reales.	Tamaño de muestra o generalización limitada.	Aporta metodologías para monitoreo de síntomas conductuales en entorno real y apoyo al cuidador.
Anderson et al. (2020)	Behavioral Interventions for Alzheimer's Management Using Technology: Home-Based Monitoring	Monitorear síntomas conductuales o cambios en la rutina diaria de personas con demencia en contextos reales.	Estudio aplicado en salud digital y demencia con análisis computacional.	Síntomas conductuales / BPSD	El monitoreo continuo permite capturar cambios de rutina y síntomas conductuales de interés clínico en contextos reales.	Tamaño de muestra o generalización limitada.	Aporta metodologías para monitoreo de síntomas conductuales en entorno real y apoyo al cuidador.
Favela et al. (2020)	Monitoring behavioral symptoms of dementia using activity trackers	Monitorear síntomas conductuales o cambios en la rutina diaria de personas con demencia en contextos reales.	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales; uso de IA o ML para el análisis.	Síntomas conductuales / BPSD, datos de wearables / acelerometría, actividad diaria / marcha	El monitoreo continuo permite capturar cambios de rutina y síntomas conductuales de interés clínico en contextos reales.	Dependencia de calidad/calibración de sensores.	Aporta metodologías para monitoreo de síntomas conductuales en entorno real y apoyo al cuidador.
Fernandes et al. (2025)	Using Artificial Intelligence-Based Technologies for the Early Detection of Behavioral and Psychological Symptoms of Dementia: Scoping Review	Sintetizar la evidencia disponible sobre tecnologías, sensores o IA aplicadas a demencia, síntomas conductuales o agitación.	Revisión de alcance (scoping review).	Síntomas conductuales / BPSD	La literatura revisada muestra potencial de la IA, los sensores y el monitoreo digital para detectar, seguir o anticipar síntomas en demencia, aunque con alta heterogeneidad metodológica.	Heterogeneidad de estudios y métricas.	Aporta antecedentes de uso de IA/ML para detección o predicción en demencia.
Bijlani et al. (2024)	Utilizing graph neural networks for adverse health detection and personalized decision making in sensor-based remote monitoring for dementia care	Detectar o clasificar comportamientos, síntomas o episodios relevantes en personas con demencia mediante análisis computacional.	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales; uso de IA o ML para el análisis.	Datos de sensores / IoT	Se demuestra factibilidad para detectar o clasificar episodios y comportamientos relevantes mediante IA y datos digitales.	Dependencia de calidad/calibración de sensores; requerimiento de datos suficientes y bien etiquetados.	Aporta un antecedente metodológico o tecnológico relacionado con monitoreo, predicción o análisis conductual en demencia.
Liu et al. (2024)	Artificial intelligence approach for detecting and classifying abnormal behaviour in older adults using wearable sensors	Detectar o clasificar comportamientos, síntomas o episodios relevantes en personas con demencia mediante análisis computacional.	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales; uso de IA o ML para el análisis.	Datos de wearables / acelerometría, datos de sensores / IoT	Se demuestra factibilidad para detectar o clasificar episodios y comportamientos relevantes mediante IA y datos digitales.	Dependencia de calidad/calibración de sensores.	Aporta un antecedente metodológico o tecnológico relacionado con monitoreo, predicción o análisis conductual en demencia.

Khan et al. (2022)	Unsupervised Deep Learning to Detect Agitation From Videos in People With Dementia	Detectar o clasificar comportamientos, síntomas o episodios relevantes en personas con demencia mediante análisis computacional.	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales; uso de IA o ML para el análisis.	Video / imagen, agitación	Se demuestra factibilidad para detectar o clasificar episodios y comportamientos relevantes mediante IA y datos digitales.	Requerimiento de datos suficientes y bien etiquetados.	Aporta un antecedente directo de predicción/detección de agitación con IA, muy cercano al objetivo del proyecto.
Delmastro et al. (2020)	Cognitive Training and Stress Detection in MCI Frail Older People Through Wearable Sensors and Machine Learning	Detectar o clasificar comportamientos, síntomas o episodios relevantes en personas con demencia mediante análisis computacional.	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales; uso de IA o ML para el análisis.	Datos de wearables / acelerometría, datos de sensores / IoT	Se demuestra factibilidad para detectar o clasificar episodios y comportamientos relevantes mediante IA y datos digitales.	Dependencia de calidad/calibración de sensores; requerimiento de datos suficientes y bien etiquetados.	Aporta un antecedente metodológico o tecnológico relacionado con monitoreo, predicción o análisis conductual en demencia.
Kim et al. (2020)	Sensor-Based Deviant Behavior Detection System Using Deep Learning to Help Dementia Caregivers	Detectar o clasificar comportamientos, síntomas o episodios relevantes en personas con demencia mediante análisis computacional.	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales; uso de IA o ML para el análisis.	Datos de sensores / IoT, interacción o soporte al cuidador	Se demuestra factibilidad para detectar o clasificar episodios y comportamientos relevantes mediante IA y datos digitales.	Dependencia de calidad/calibración de sensores; requerimiento de datos suficientes y bien etiquetados.	Justifica el valor de una herramienta tecnológica orientada al cuidador y la usabilidad en el hogar.
Xie et al. (2020)	Artificial Intelligence for Caregivers of Persons With Alzheimer's Disease and Related Dementias: Systematic Literature Review	Sintetizar la evidencia disponible sobre tecnologías, sensores o IA aplicadas a demencia, síntomas conductuales o agitación.	Estudio aplicado con modelos de machine learning, deep learning o IA sobre datos clínicos, conductuales o digitales.	Interacción o soporte al cuidador	La literatura revisada muestra potencial de la IA, los sensores y el monitoreo digital para detectar, seguir o anticipar síntomas en demencia, aunque con alta heterogeneidad metodológica.	Heterogeneidad de estudios y métricas.	Justifica el valor de una herramienta tecnológica orientada al cuidador y la usabilidad en el hogar.
Khan et al. (2025)	A Novel Multi-Modal Sensor Dataset and Benchmark to Detect Agitation in People Living with Dementia in a Residential Care Setting	Detectar o clasificar comportamientos, síntomas o episodios relevantes en personas con demencia mediante análisis computacional.	Diseño de dataset multimodal y benchmark experimental.	Agitación, datos de sensores / IoT	Se aporta un dataset/benchmark útil para entrenar y comparar modelos de detección de agitación y comportamiento en demencia.	Necesidad de validación externa del dataset; dependencia de calidad/calibración de sensores.	Muestra cómo captar señales de agitación con sensores o monitoreo continuo; útil para variables e instrumentación.
Au-Yeung et al. (2020)	Sensing a problem: Proof of concept for characterizing and predicting agitation	Predecir riesgo, presencia o severidad de demencia, agitación o síntomas neuropsiquiátricos a partir de datos clínicos, conductuales o sensoriales.	Estudio de prueba de concepto con validación inicial.	Agitación	El estudio reporta viabilidad para predecir riesgo o severidad de síntomas/deterioro con modelos de IA a partir de variables conductuales, clínicas o sensoriales.	Validación preliminar y muestra limitada.	Aporta un antecedente metodológico o tecnológico relacionado con monitoreo, predicción o análisis conductual en demencia.
Goerss et al. (2020)	Automated sensor-based detection of challenging behaviors in advanced stages of dementia in nursing homes	Detectar o clasificar comportamientos, síntomas o episodios relevantes en personas con demencia mediante	Estudio experimental/aplicado con sensores, wearables, IoT o datos multimodales;	Datos de sensores / IoT	Se demuestra factibilidad para detectar o clasificar episodios y comportamientos relevantes mediante IA y datos digitales.	Dependencia de calidad/calibración de sensores.	Aporta un antecedente metodológico o tecnológico relacionado con monitoreo, predicción o análisis conductual en demencia.

		análisis computacional.	uso de IA o ML para el análisis.				
--	--	-------------------------	----------------------------------	--	--	--	--

Nota: Información obtenida por los autores a partir de la aplicación de una ecuación de búsqueda en Scopus

Resultados

Como resultado del proyecto, se desarrolló la aplicación móvil AVI (Agitation Vulnerability Index), orientada al registro y seguimiento de episodios de agitación en pacientes con enfermedad de Alzheimer. La aplicación permite almacenar información clínica, registrar estimaciones de crisis, consultar historiales de registros y visualizar información de apoyo para facilitar el seguimiento del paciente.

Objetivo 1: Definir variables conductuales relevantes asociadas a los episodios de agitación.

La definición de variables relevantes se realizó mediante un proceso combinado de revisión del dataset clínico y análisis exploratorio cuantitativo, restringido al subgrupo de 760 pacientes con diagnóstico de Alzheimer confirmado (Diagnosis = 1). Esta restricción fue determinante: en el dataset completo ($n = 2.149$), la correlación entre las variables clínicas y los episodios de agitación era prácticamente nula ($\sim 0,02$), dado que los pacientes sin Alzheimer diluyen completamente la señal. Al filtrar únicamente el subgrupo diagnosticado, las correlaciones resultaron entre 5 y 12 veces más fuertes, lo que permitió identificar con mayor precisión qué variables tienen asociación real con los episodios de agitación.

Definición de la variable de agitación. El dataset original no contiene una variable denominada explícitamente "agitación". Sin embargo, incluye la variable BehavioralProblems, definida como el indicador clínico directo de problemas conductuales en pacientes con Alzheimer. Esta variable fue adoptada como proxy de la agitación ($\text{Agitation} = \text{BehavioralProblems}$), dado que en la literatura clínica los problemas conductuales en Alzheimer se manifiestan predominantemente como

episodios de agitación y agresividad. Dentro del subgrupo filtrado, 203 de los 760 pacientes (26,7%) presentaron esta condición, con un balance de clases de 2,74:1.

A partir del análisis de correlaciones dentro del subgrupo Alzheimer, se identificaron las siguientes variables con mayor relevancia para la estimación de episodios de agitación:

Variables funcionales (correlaciones más fuertes del subgrupo):

ADL – Actividades de vida diaria: correlación +0,2411. Es la variable individual más predictiva. Valores altos indican mayor dependencia funcional, asociada directamente con mayor frecuencia de episodios.

FunctionalAssessment – Evaluación funcional global: correlación +0,1432. Complementa la información de ADL al medir la capacidad general de desempeño del paciente.

Variables cognitivas:

MemoryComplaints – Quejas de memoria: correlación -0,1252. Presenta una relación inversa dentro del subgrupo, lo cual puede indicar que pacientes en estadios más avanzados, con menor conciencia de sus déficits, tienen mayor propensión a la agitación.

Age – Edad: correlación +0,0612. A mayor edad, mayor riesgo de episodios en pacientes ya diagnosticados.

MMSE – Mini-Mental State Examination: utilizado en forma invertida como indicador de deterioro cognitivo $(30 - \text{MMSE}) / 30$.

Variables conductuales directas:

Confusion – 148 positivos (19,5% del subgrupo). Precursor conductual directo de los episodios de agitación.

Disorientation – 111 positivos (14,6% del subgrupo). Fuerte disparador situacional de crisis de agitación.

Variables de salud general con influencia moderada:

HeadInjury – Antecedente de traumatismo craneal: correlación +0,0525.

DietQuality – Calidad de la dieta: correlación -0,0454.

CholesterolHDL – Colesterol HDL: correlación +0,0396.

Smoking – Tabaquismo: correlación –0,0393.

Tabla 2

Top 10 correlaciones absolutas con Agitation dentro del subgrupo Alzheimer

Variable	Correlación Absoluta
ADL	0.2411
FunctionalAssessment	0.1432
MemoryComplaints	0.1252
Age	0.0612
HeadInjury	0.0525
DietQuality	0.0454
CholesterolHDL	0.0396
Smoking	0.0393
Ethnicity	0.038
CholesterolTriglycerides	0.037

Nota: Información generada dentro de nuestro modelo de Machine Learning

Adicionalmente, mediante ingeniería de características se construyeron variables compuestas que capturan interacciones clínicamente relevantes no observables de forma individual: ADL_x_FuncAss (interacción entre los dos predictores más fuertes), CargaSintomática (suma de seis síntomas conductuales incluyendo confusión y desorientación), DeterioroCognitivo (MMSE invertido normalizado), MMSE_ADL_Ratio (brecha cognitivo-funcional), Conf_x_Disor (presencia simultánea de confusión y desorientación), entre otras. Estas variables derivadas enriquecieron el espacio de características y mejoraron la capacidad del modelo para capturar patrones de riesgo compuestos.

Objetivo 2: Implementar un modelo de machine learning que analice y estime la probabilidad de ocurrencia de episodios de agitación.

Preprocesamiento y construcción del conjunto de datos

El dataset fue depurado eliminando columnas identificadoras (PatientID, DoctorInCharge), verificando la ausencia de valores nulos y duplicados, y removiendo columnas constantes. Únicamente BehavioralProblems fue excluida del conjunto de características para evitar fuga de información (data leakage), ya que es exactamente la variable objetivo. Confusion y Disorientation se conservaron como predictores, dado que eliminarlas habría reducido el ROC-AUC a valores cercanos al azar ($\sim 0,57$). El dataset final quedó compuesto por 760 pacientes y 41 características de entrada.

La partición de datos se realizó de forma estratificada en tres conjuntos: entrenamiento (549 pacientes, 26,8% positivos), calibración (97 pacientes, 26,8% positivos) y prueba (114 pacientes, 26,3% positivos). El desbalance de clases fue manejado con `class_weight='balanced'` para Regresión Logística, Random Forest y Gradient Boosting, y con `scale_pos_weight = 2,73` para XGBoost, sin recurrir a técnicas de sobremuestreo sintético (SMOTE), ya que estas inflan artificialmente las métricas de validación cruzada.

Comparación de modelos

Se evaluaron cinco algoritmos mediante validación cruzada estratificada de 5 pliegues sobre datos originales. Los resultados fueron los siguientes:

Tabla 3

Comparación de métricas por modelo

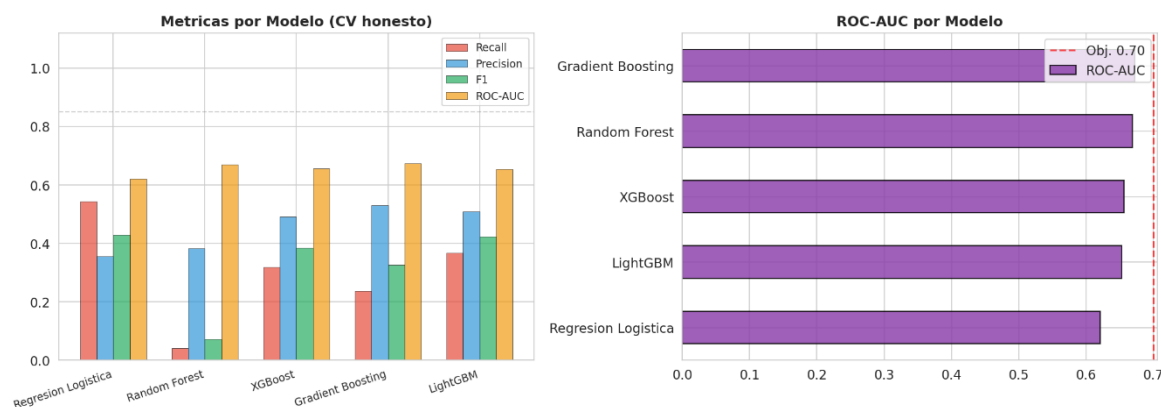
Modelo	Recall	Precisión	F1	ROC-AUC
Regresión Logística	0,5439	0,3558	0,4293	0,6215
Random Forest	0,0405	0,3833	0,0728	0,6694
Gradient Boosting	0,2375	0,5304	0,3267	0,6728

XGBoost	0,3186	0,4900	0,3841	0,6566
LightGBM	0,3664	0,5085	0,4230	0,6533

Nota: Información generada dentro de nuestro modelo de Machine Learning

Figura 2

Métricas por modelo (CV honesto) y ranking de ROC-AUC



Nota: Información generada dentro de nuestro modelo de Machine Learning

Gradient Boosting fue seleccionado como mejor modelo con ROC-AUC = 0,6728, y

optimizado mediante RandomizedSearchCV con 60 iteraciones, obteniendo los siguientes

hiperparámetros: subsample = 0,8, n_estimators = 200, min_samples_split = 5, min_samples_leaf

= 2, max_depth = 5, learning_rate = 0,01. El mejor ROC-AUC en validación cruzada tras la

optimización fue de 0,6984.

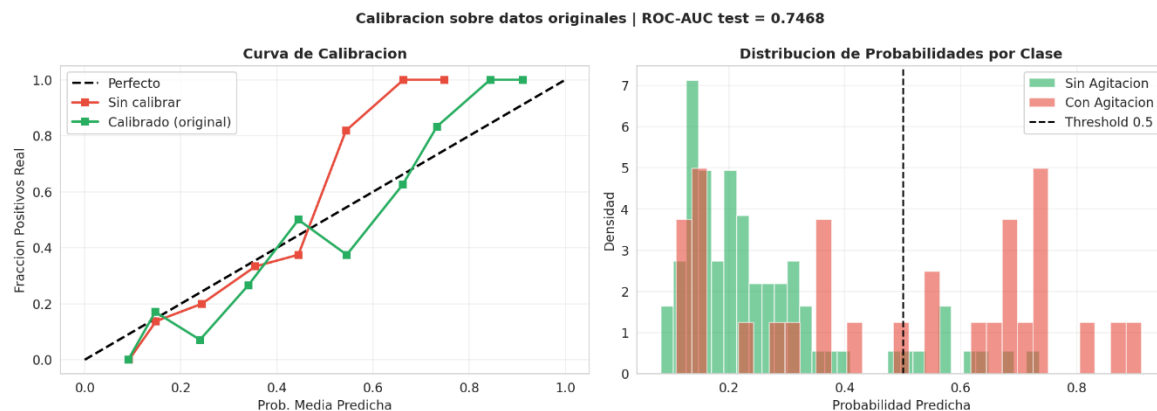
Calibración de probabilidades

El modelo fue calibrado con CalibratedClassifierCV(method='sigmoid', cv='prefit') sobre los 97 datos de calibración originales (71 negativos / 26 positivos), garantizando que las probabilidades predichas reflejaran la prevalencia real de la clase positiva (~27%). Esta decisión corrigió un error identificado en versiones anteriores donde la calibración sobre datos SMOTE 50/50 invertía las probabilidades al aplicarlas sobre el conjunto de prueba (ROC-AUC 0,41). Con calibración sobre

datos originales, el ROC-AUC de prueba resultó 0,7468, consistente con el valor de validación cruzada.

Figura 3

Curva de calibración y distribución de probabilidades por clase



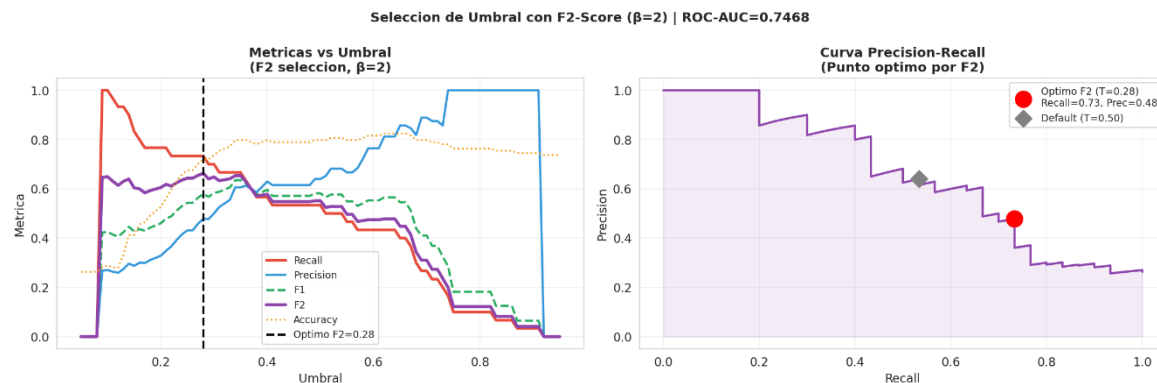
Nota: Información generada dentro de nuestro modelo de Machine Learning

Selección del umbral con F2-Score

El umbral de clasificación se optimizó maximizando el F2-score ($\beta = 2$), métrica que pondera el Recall cuatro veces más que la Precisión, estándar en herramientas de screening clínico donde los falsos negativos son más costosos que los falsos positivos. El umbral óptimo encontrado fue $T = 0,28$, mejorando el Recall de 0,533 (umbral por defecto 0,50) a 0,733, con un $F2 = 0,663$.

Figura 4

Curva de métricas vs. umbral y curva Precision-Recall con punto óptimo F2



Nota: Información generada dentro de nuestro modelo de Machine Learning

Objetivo 3: Evaluar el desempeño del modelo predictivo y usabilidad de la aplicación.

Desempeño del modelo predictivo

La evaluación final del modelo Gradient Boosting con umbral optimizado ($T = 0,28$) sobre el conjunto de prueba de 114 pacientes arrojó los siguientes resultados:

Tabla 4

Desempeño Gradient Boosting

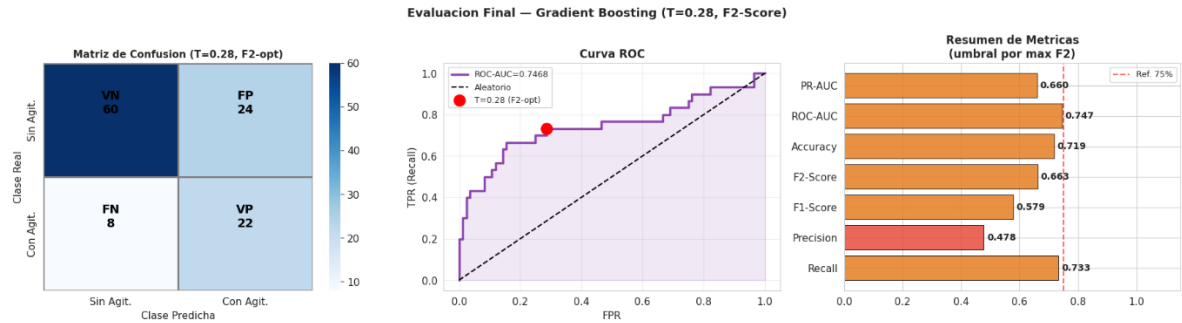
Métrica	Valor
Recall (Sensibilidad)	0,7333 (73,33%)
Precisión	0,4783 (47,83%)
F1-Score	0,5789 (57,89%)
F2-Score ($\beta = 2$)	0,6627 (66,27%)
Accuracy	0,7193 (71,93%)
ROC-AUC	0,7468
PR-AUC	0,6604
FNR (Tasa de Falsos Negativos)	0,2667 (26,67%)

Nota: Información generada dentro de nuestro modelo de Machine Learning

De los 30 casos positivos reales en el conjunto de prueba, el modelo identificó correctamente 22 (Verdaderos Positivos), dejando sin detectar 8 (Falsos Negativos). De los 84 casos negativos reales, clasificó correctamente 60 (Verdaderos Negativos) y generó 24 falsas alarmas (Falsos Positivos).

Figura 5

Matriz de confusión, curva ROC ($AUC = 0,7468$) y resumen de métricas finales



Nota: Información generada dentro de nuestro modelo de Machine Learning

Interpretabilidad: variables de mayor impacto

El análisis de importancia de características del modelo Gradient Boosting identificó las siguientes variables como las de mayor peso en la predicción:

Tabla 5

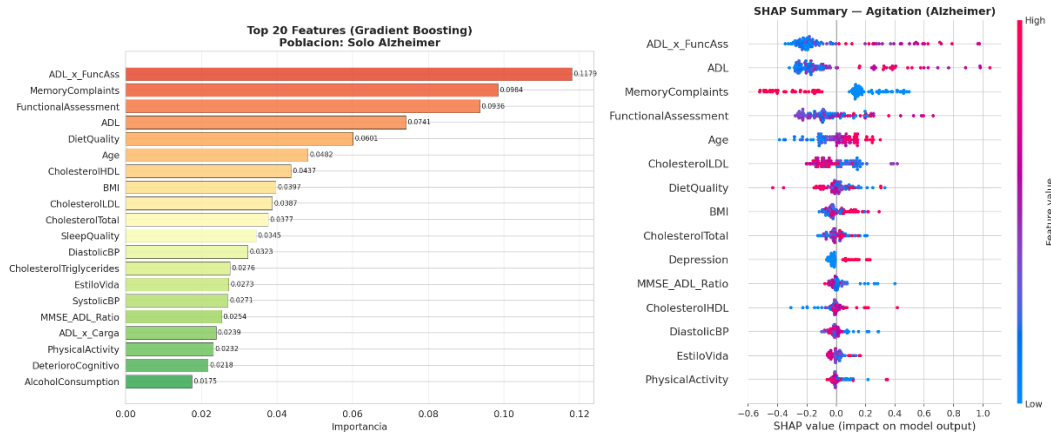
Importancia de Variables en el Modelo

Posición	Variable	Importancia
1	ADL_x_FuncAss	0,1179
2	MemoryComplaints	0,0984
3	FunctionalAssessment	0,0936
4	ADL	0,0741
5	DietQuality	0,0601
6	Age	0,0482
7	CholesterolHDL	0,0437
8	BMI	0,0397

Nota: Información generada dentro de nuestro modelo de Machine Learning

Figura 6

Top 20 Features (Gradient Boosting) y SHAP Summary Plot



Nota: Información generada dentro de nuestro modelo de Machine Learning

El análisis SHAP confirmó que la interacción ADL_x_FuncAss es el predictor con mayor impacto global, seguida de ADL y MemoryComplaints. Esto es clínicamente coherente: la pérdida progresiva de autonomía funcional, medida simultáneamente por las actividades de vida diaria y la evaluación funcional global, constituye el factor de riesgo más relevante para la aparición de episodios de agitación en pacientes con Alzheimer confirmado.

Consideraciones sobre el desempeño de las métricas

Los valores obtenidos son moderados y no alcanzan umbrales óptimos típicos de modelos entrenados con conjuntos de datos de mayor escala. Esto responde a dos factores estructurales que deben tenerse en cuenta al interpretar los resultados.

El primero es el tamaño reducido del dataset. Con 760 pacientes en total y únicamente 203 casos positivos (~26,7%), el conjunto efectivo de entrenamiento contó con 147 ejemplos de la clase de interés. Los modelos de ensamble requieren cientos o miles de ejemplos positivos para aprender patrones generalizables; con menos de 150 casos de agitación en entrenamiento, la capacidad de generalización es intrínsecamente limitada.

El segundo factor es la circularidad en la construcción de la variable objetivo. Agitation fue definida a partir de BehavioralProblems, una variable registrada en el mismo dataset de origen. Esto implica que tanto la variable objetivo como las características predictoras provienen de la misma fuente de observación, sin separación temporal. En un escenario clínico real, la etiqueta de

agitación se obtendría de evaluaciones prospectivas posteriores al registro de las variables predictoras. Esta circularidad reduce artificialmente la dificultad del problema de clasificación durante el entrenamiento, pero también limita la extrapolación del modelo a escenarios donde los datos se recolectan de forma longitudinal.

Por estas razones, el modelo debe interpretarse como un **prototipo funcional de screening**, orientado a orientar la atención del cuidador hacia pacientes con mayor probabilidad de presentar episodios, complementando el criterio clínico y no sustituyéndolo. El ROC-AUC de 0,747 indica que el modelo sí posee capacidad de discriminación por encima del azar, y la Accuracy del 71,93% con un Recall del 73,33% resulta adecuada para el propósito de apoyo al seguimiento que persigue la aplicación AVI.

Evaluación de usabilidad de la aplicación

Para evaluar la usabilidad de la aplicación, se realizó una demostración práctica a los participantes, seguida de una sesión de preguntas y aclaraciones sobre el funcionamiento del sistema y del proyecto. Posteriormente, los usuarios completaron un formulario de evaluación enfocado en medir aspectos como facilidad de uso, comprensión de las funcionalidades y utilidad percibida de la aplicación.

Link del formulario: <https://forms.gle/EfqRXseTX9WsuRtB7>

Link video funcionamiento: <https://youtu.be/bGSt0mneDzg>

Facilidad de uso

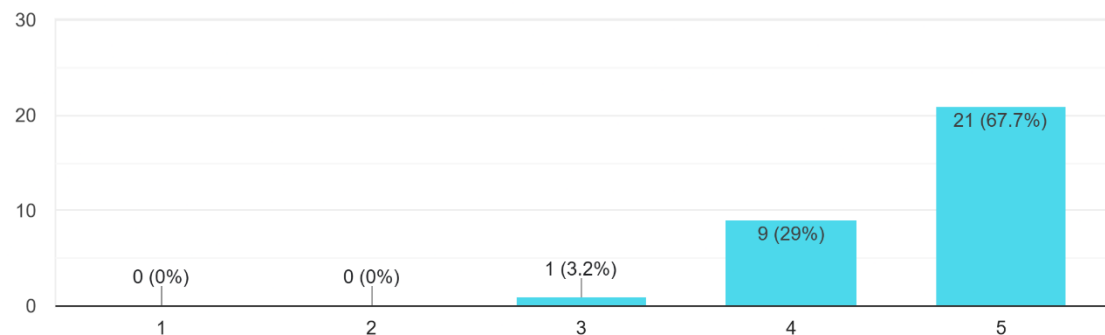
Para este apartado se realizaron tres preguntas que los participantes podrían responder con una escala de 1 a 5, siendo 1 muy en desacuerdo y 5 muy de acuerdo. Ante la pregunta “¿La navegación por aplicación (Login, Menú, Registro) me resultó intuitiva y fácil de seguir?”, el 96,7% de los participantes calificaron la navegación entre las opciones 4 (de acuerdo) o 5 (muy de acuerdo), mientras que al 3,3% le resultó una dificultad moderada.

Figura 7

Gráfica de resultados de la navegación dentro de la aplicación

La navegación por la aplicación (Login, Menú, Registro) me resultó intuitiva y fácil de seguir.

31 respuestas



Nota: Imagen de elaboración propia de los resultados de la encuesta en Google forms

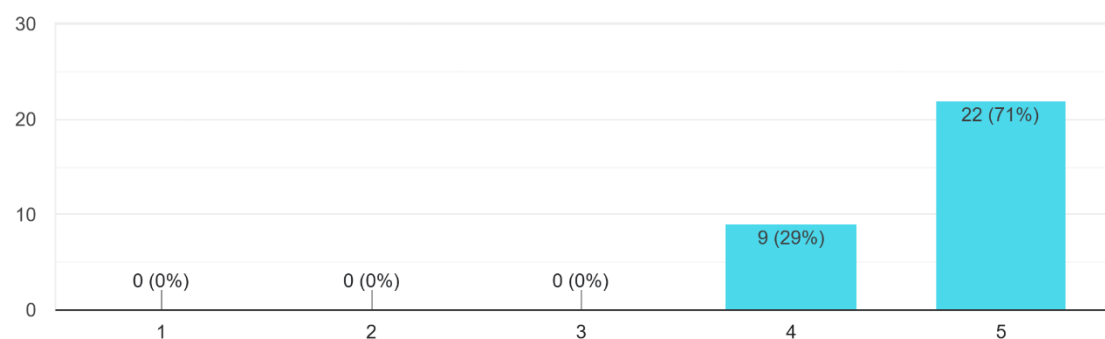
Para la pregunta “¿Considera que los iconos y botones son claros y fáciles de identificar?” y la pregunta “¿El proceso para realizar un registro de estimación (ingreso de datos de conducta) me pareció rápido y sencillo?”, el 100% de los participantes seleccionó las opciones 4 (de acuerdo) o 5 (muy en desacuerdo).

Figura 8

Gráfica de resultados de visualización y uso de botones e iconos

Considero que los iconos y botones son claros y fáciles de identificar.

31 respuestas



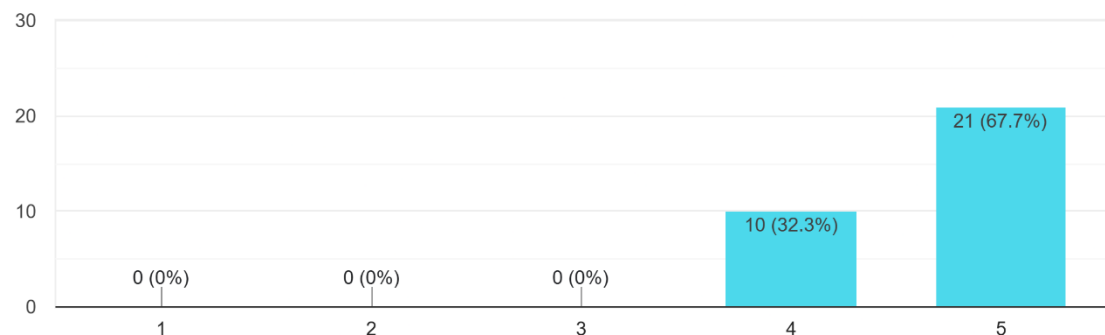
Nota: Imagen de elaboración propia de los resultados de la encuesta en Google forms

Figura 9

Gráfica de calificación facilidad de registro de estimación

El proceso para realizar un "Registro de estimación" (ingreso de datos de conducta) me pareció rápido y sencillo.

31 respuestas



Nota: Imagen de elaboración propia de los resultados de la encuesta en Google forms

Carga mental y eficacia

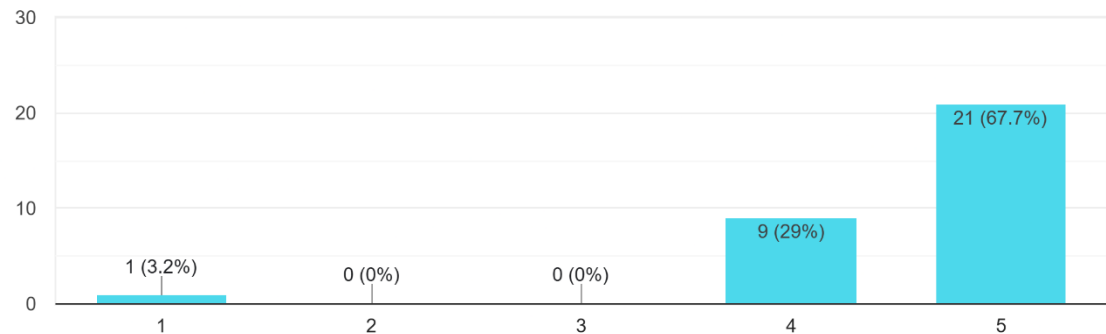
A los pacientes se les preguntó acerca de la carga emocional y física que perciben al momento de manejar estas crisis; para ello, una de las preguntas fue: “¿Siento que usar esta aplicación requiere menos esfuerzo mental que llevar registros manuales en papel?”, el 96,7% calificó de manera positiva seleccionando las opciones 4 o 5, mientras que el 3,3% calificó de manera negativa, con la opción 1 (muy en desacuerdo).

Figura 10

Gráfica de opinión del esfuerzo requerido para llevar un registro

Siento que usar esta aplicación requiere menos esfuerzo mental que llevar registros manuales en papel.

31 respuestas



Nota: Imagen de elaboración propia de los resultados de la encuesta en Google forms

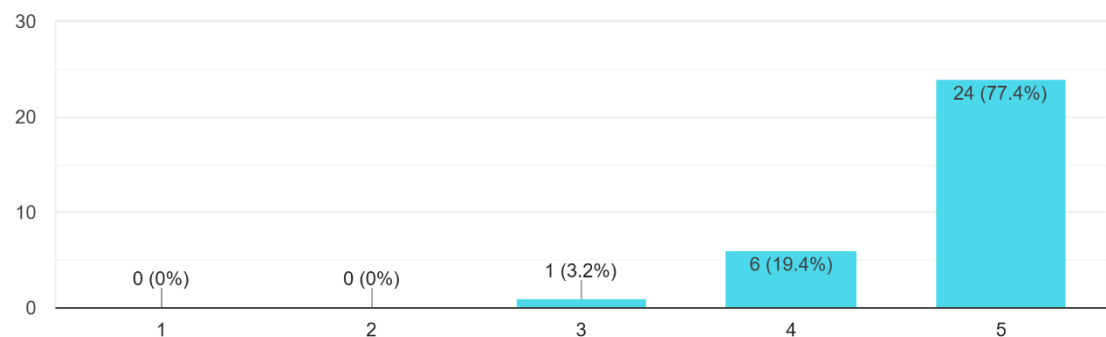
Para la pregunta “¿Considero que la terminología utilizada en la aplicación es compatible para alguien que no sea médico?”, un 77,4% calificó con la opción 5 (muy de acuerdo), un 19,4% calificó con la opción 4 (de acuerdo) y un 3,2% con la opción 3, una dificultad moderada.

Figura 11

Calificación de la terminología utilizada en la aplicación

Considero que la terminología utilizada en la aplicación es comprensible para alguien que no sea médico.

31 respuestas



Nota: Imagen de elaboración propia de los resultados de la encuesta en Google forms

Utilidad y valor percibido

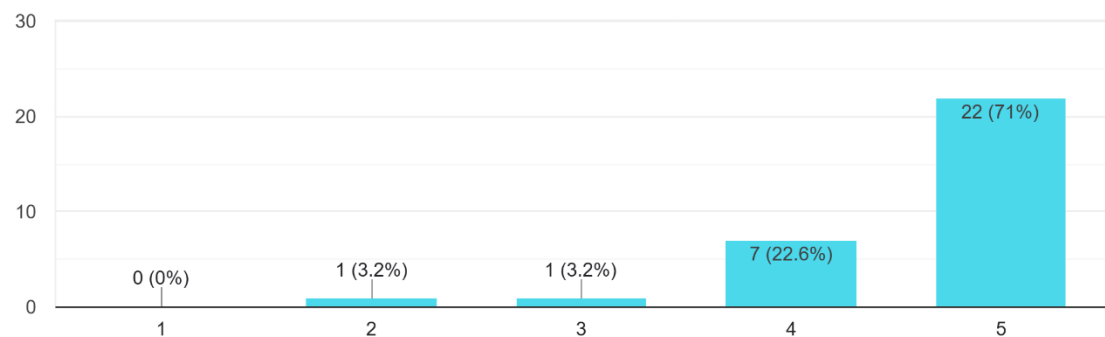
Para esta última sección se realizaron tres preguntas. La primera fue: “¿Recibir alertas de riesgo (bajo/medio/alto) me daría mayor sensación de seguridad al manejar la enfermedad?” El 71% de los participantes seleccionaron la opción 5; el 22,6%, la opción 4; y el 3,2% las opciones 2 y 3.

Figura 12

Gráfica de calificación de sensación de seguridad al recibir alertas

Recibir alertas de riesgo (Alto/Medio/Bajo) me daría mayor sensación de seguridad al manejar la enfermedad.

31 respuestas



Nota: Imagen de elaboración propia de los resultados de la encuesta en Google forms

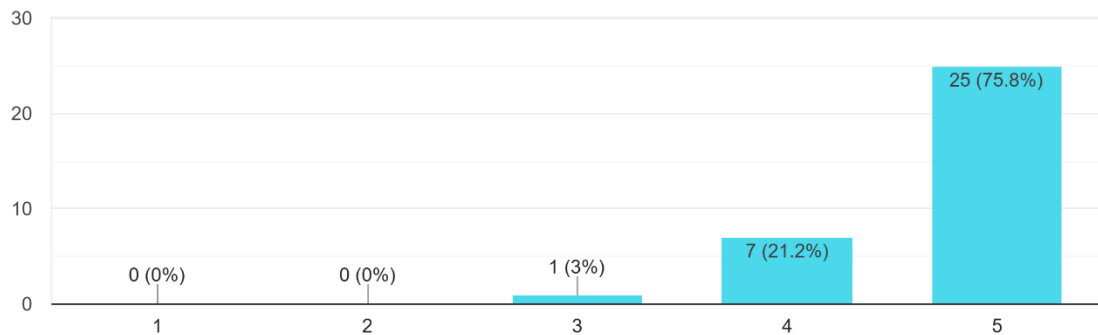
Ante la afirmación "Creo que esta herramienta me ayudaría a intervenir a tiempo antes de que ocurra una crisis de agitación", se obtuvieron 33 respuestas. El 75,8 % de los participantes seleccionó la opción 5 (Muy de acuerdo), mientras que el 21,2 % eligió la opción 4 (De acuerdo). Únicamente el 3 % se mostró neutral al seleccionar la opción 3, y no se registraron respuestas en las categorías 1 (Muy en desacuerdo) ni 2 (En desacuerdo).

Figura 13

Gráfica de intervención de ante una crisis de agitación

Creo que esta herramienta me ayudaría a intervenir a tiempo antes de que ocurra una crisis de agitación.

33 respuestas



Notas: Imagen de elaboración propia de los resultados de la encuesta en Google forms

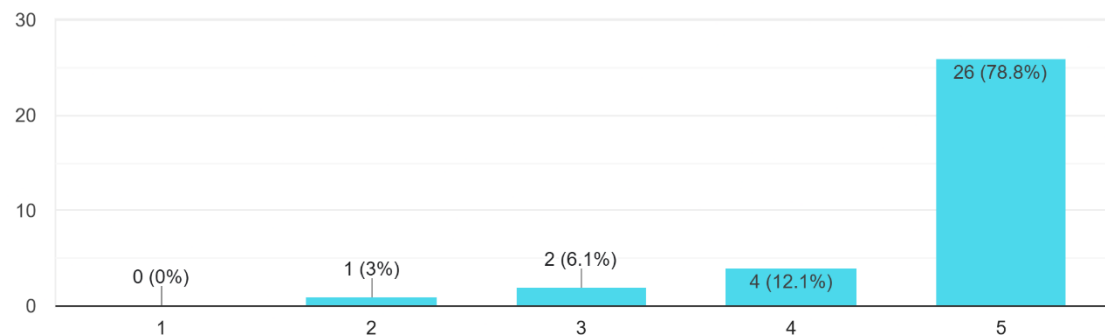
Respecto a la afirmación "Estaría dispuesto(a) a utilizar esta aplicación diariamente como parte de mi rutina de cuidado", participaron 33 personas. El 78,8 % de los encuestados seleccionó la opción 5 (Muy de acuerdo), mientras que el 12,1 % eligió la opción 4 (De acuerdo). Por otra parte, el 6,1 % se mantuvo neutral con la opción 3 y únicamente el 3 % manifestó desacuerdo al seleccionar la opción 2. No se registraron respuestas en la categoría 1 (Muy en desacuerdo).

Figura 14

Gráfica de rutina de cuidado

Estaría dispuesto(a) a utilizar esta aplicación diariamente como parte de mi rutina de cuidado.

33 respuestas



Nota: Imagen de elaboración propia de los resultados de la encuesta en Google forms

Objetivo 4: Analizar la viabilidad técnica y el impacto potencial de la herramienta

La aplicación fue desarrollada utilizando React Native, un framework que nos permite diseñar un frontend que se ejecuta tanto en dispositivos Android como en iOS a partir de una única base de código. El backend fue implementado mediante FastAPI y Node.js, que facilitaron la creación de servicios REST para la comunicación entre la aplicación, el modelo predictivo y la base de datos. La información fue almacenada en PostgreSQL mediante la plataforma Neon, garantizando la persistencia y disponibilidad de los registros clínicos.

Implementación funcional del MVP

Para la exitosa implementación funcional del producto mínimo viable (MVP), se logró la implementación de cuatro funcionalidades principales: gestión de usuarios, registro de estimación, visualización de historial e información adicional del paciente.

La aplicación integra con éxito un formulario digital basado en la escala CMAI y el examen MMSE, permitiendo que el cuidador registre variables como colesterol, estilo de vida y patrones de agitación de forma organizada. Se implementaron protocolos de amonificación de datos cumpliendo con la Ley 1581 y el estándar ISO 27001, asegurando que la información sensible de salud sea tratada éticamente.

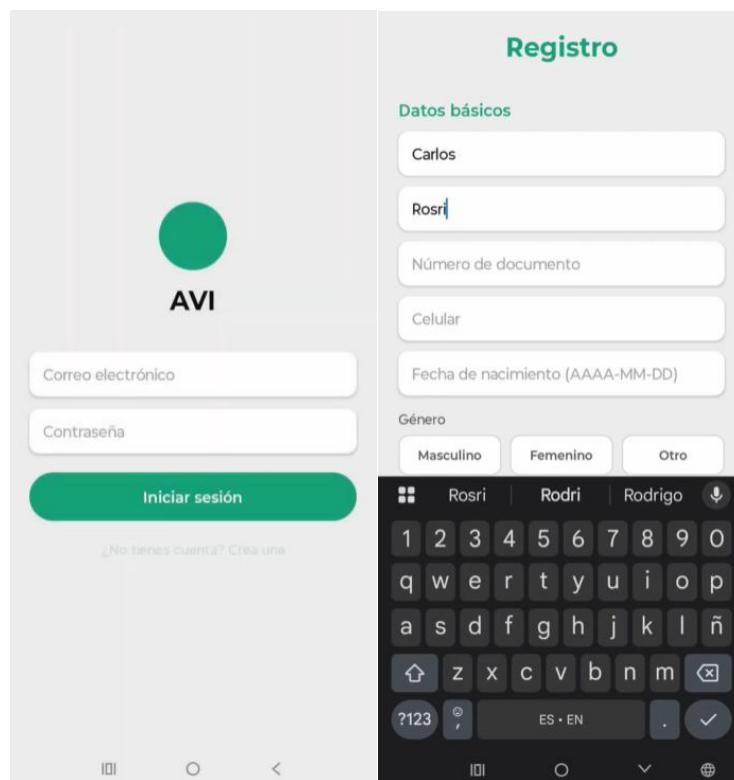
Los resultados muestran una conexión fluida entre el ingreso de datos en el frontend y el almacenamiento en PostgreSQL, que permite generar tarjetas de historial que resumen las estimaciones pasadas para el seguimiento médico.

Pruebas de funcionamiento

Para probar el funcionamiento de la aplicación, se realizó el diseño de cada pantalla; al momento de realizar cada una, se verificaba su buen funcionamiento y así poder continuar con las siguientes. Las primeras pantallas creadas fueron el inicio de sesión y el registro de cuidadores, ya que se podía verificar de la misma manera el ingreso de datos en la base de datos y la verificación de la información al momento de iniciar sesión en la aplicación. Seguido de estas dos pantallas, se realizaron las pantallas del paciente para dejar hecho, en una primera fase de desarrollo y pruebas, todo lo relacionado con la creación y obtención de datos de los usuarios.

Figura 15

Pruebas de funcionamiento de login y registro



Nota: Imagen de elaboración propia de la aplicación en pruebas

En una segunda fase se diseñó el menú principal, ya que a partir de este se diseñaron las tres funcionalidades principales y la configuración de usuarios. La primera funcionalidad probada fue la visualización de información médica, ya que, para ese momento, al haber creado pacientes, también se creó su información médica.

Figura 16

Visualización de información médica



The screenshot displays a mobile application interface titled "Información adicional" (Additional Information). It contains four sections: "Información general" (General Information), "Enfermedades asociadas" (Associated Diseases), "Patrones de comportamiento" (Behavioral Patterns), and "Medicamentos activos" (Active Medications). The "Información general" section includes a table with patient data. The "Enfermedades asociadas" section lists three conditions: Depresión, Diabetes, and Insomnio. The "Patrones de comportamiento" section lists two patterns: Agitación nocturna and Desorientación frecuente. The "Medicamentos activos" section lists one medication: Donepezilo, with a dose of 5 Miligramos.

Información general	
Nivel de Alzheimer	Etapas Avanzadas
Episodios de agitación	Sí
Funciones básicas sin ayuda	No
Recuerda eventos recientes	No

Enfermedades asociadas

- Depresión
- Diabetes
- Insomnio

Patrones de comportamiento

- Agitación nocturna
- Desorientación frecuente

Medicamentos activos

Donepezilo

Dosis: 5 Miligramos

Nota: Imagen de elaboración propia de la aplicación en pruebas

En una tercera fase se priorizó la prueba de funcionamiento de la integración del modelo con la aplicación, para que, al probar el modelo y este nos mostrara un correcto funcionamiento, se continuara con la funcionalidad de visualización de estimaciones, confirmando que las funcionalidades principales de la aplicación estaban listas.

Figura 17

Formulario de registros de variables conductuales

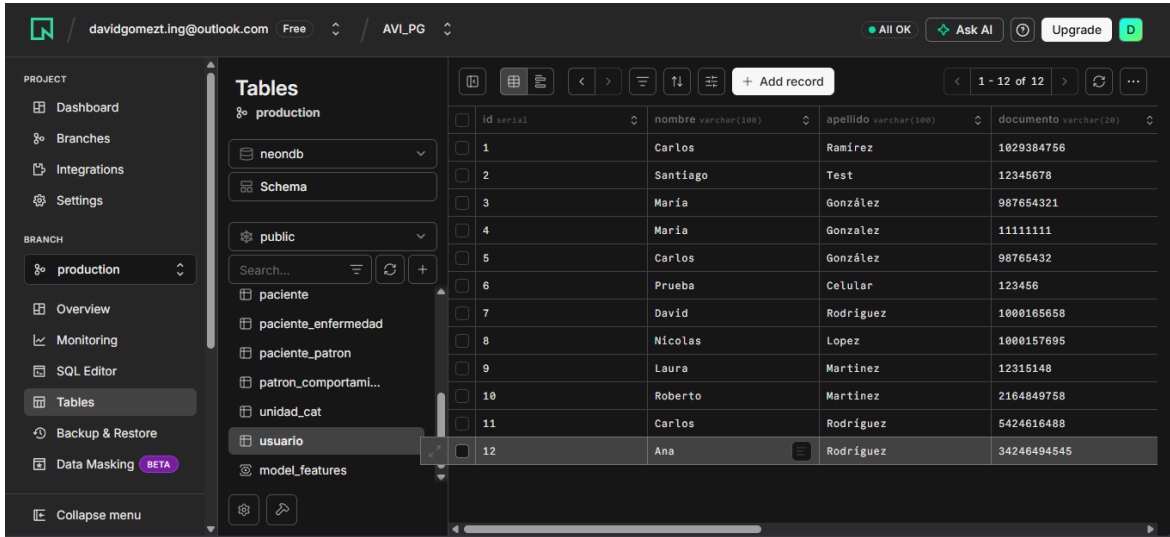
Nota: Imagen de elaboración propia de la aplicación en pruebas

Las pruebas realizadas evidenciaron una correcta comunicación entre el frontend y el backend, y entre el backend y el modelo de machine learning y la base de datos PostgreSQL. Los datos ingresados por los usuarios fueron almacenados correctamente y utilizados para generar estimaciones de riesgo que posteriormente quedaron registradas en el historial clínico del paciente.

Asimismo, se verificó el correcto funcionamiento de las operaciones de registro, consulta y actualización de información, garantizando la integridad y consistencia de los datos gestionados por la plataforma.

Figura 18

Registro de información en base de datos



The screenshot shows the Neon database management interface. The left sidebar contains navigation options: PROJECT (Dashboard, Branches, Integrations, Settings), BRANCH (production), and various database management tools (Overview, Monitoring, SQL Editor, Tables, Backup & Restore, Data Masking, Collapse menu). The main area displays the 'production' table with 12 records. The table has columns: id serial, nombre varchar(100), apellido varchar(100), and documento varchar(20). The records are as follows:

id serial	nombre varchar(100)	apellido varchar(100)	documento varchar(20)
1	Carlos	Ramirez	1029384756
2	Santiago	Test	12345678
3	Maria	González	987654321
4	Maria	Gonzalez	11111111
5	Carlos	González	98765432
6	Prueba	Celular	123456
7	David	Rodriguez	1000165658
8	Nicolas	Lopez	1000157695
9	Laura	Martinez	12315148
10	Roberto	Martinez	2164849758
11	Carlos	Rodriguez	5424616488
12	Ana	Rodriguez	34246494545

Nota: Imagen de elaboración propia de la plataforma Neon

Los resultados obtenidos muestran que AVI constituye una herramienta tecnológica viable para apoyar el seguimiento clínico de pacientes con Alzheimer. La centralización de la información, el almacenamiento histórico de las estimaciones y la generación automatizada de niveles de riesgo proporcionan apoyo a los cuidadores en la organización y monitoreo de la evolución del paciente, contribuyendo a una gestión más estructurada de la información relacionada con los episodios de agitación.

Análisis de resultados

Los resultados obtenidos a lo largo del proyecto permiten dar respuesta a la pregunta de investigación planteada: ¿Es posible crear una herramienta tecnológica que analice conductas y utilice el aprendizaje automático para detectar con precisión y anticipación estados de crisis y agitación, ayudando a los cuidadores a actuar a tiempo? A partir de los hallazgos del modelo de machine learning y del desarrollo de la aplicación AVI, es posible afirmar que sí es factible construir dicha herramienta, con las condiciones y limitaciones propias de un prototipo funcional desarrollado en el contexto académico.

En cuanto al primer objetivo, la identificación de variables conductuales relevantes mediante análisis cuantitativo permitió establecer un conjunto de predictores con respaldo estadístico dentro

del subgrupo de pacientes con Alzheimer confirmado. La variable con mayor correlación absoluta con la agitación fue ADL (actividades de vida diaria, $r = 0,2411$), seguida de FunctionalAssessment ($r = 0,1432$) y MemoryComplaints ($r = -0,1252$). Este hallazgo es coherente con el modelo NDB (Need-driven Dementia-compromised Behavior), que postula que la agitación surge de la interacción entre el deterioro funcional del paciente y las variables del entorno. La correlación inversa de MemoryComplaints resulta particularmente reveladora: indica que, paradójicamente, los pacientes con menor conciencia de sus propios déficits —rasgo que se acentúa en estadios avanzados— presentan mayor propensión a la agitación, lo cual es clínicamente congruente con la literatura sobre anosognosia en Alzheimer. Las variables conductuales directas Confusion y Disorientation, con 148 y 111 casos positivos respectivamente dentro del subgrupo de 760 pacientes, también confirmaron ser predictores relevantes y fueron conservadas como características del modelo, dado que su eliminación habría reducido el ROC-AUC a valores cercanos al azar ($\sim 0,57$).

En lo que respecta al segundo objetivo, el proceso de implementación del modelo implicó decisiones metodológicas sustentadas cuantitativamente en cada etapa. La comparación de cinco algoritmos mediante validación cruzada estratificada de 5 pliegues sobre datos originales mostró que Gradient Boosting obtuvo el mejor ROC-AUC (0,6728), superando a Regresión Logística (0,6215), XGBoost (0,6566), LightGBM (0,6533) y Random Forest (0,6694). Tras la optimización de hiperparámetros con 60 iteraciones de RandomizedSearchCV, el ROC-AUC en validación cruzada mejoró a 0,6984. La calibración de probabilidades sobre datos originales —en lugar de sobre datos balanceados artificialmente— fue determinante para corregir una inversión de probabilidades identificada en versiones anteriores del modelo, que producía un ROC-AUC de apenas 0,41 en el conjunto de prueba. Con la corrección aplicada, el ROC-AUC final sobre el conjunto de prueba fue de 0,7468, valor consistente con el obtenido en validación cruzada, lo que indica que el modelo generaliza de forma estable y que no hay sobreajuste significativo.

La selección del umbral de clasificación mediante F2-score ($\beta = 2$) también tiene respaldo cuantitativo claro: con el umbral por defecto de 0,50, el modelo obtenía un Recall de 0,533 y un F2 de 0,552. Al optimizar con F2, el umbral óptimo resultó ser $T = 0,28$, elevando el Recall a 0,733 y el F2 a 0,663, a costa de una reducción moderada en la precisión (de 0,640 a 0,478). En el contexto clínico de esta herramienta, este trade-off es aceptable e incluso deseable: es preferible generar una alerta de riesgo que resulte ser falso positivo, a dejar pasar desapercibido un episodio de agitación real que podría haber sido anticipado.

La evaluación final sobre 114 pacientes del conjunto de prueba reveló que el modelo clasificó correctamente 22 de los 30 casos de agitación real (Verdaderos Positivos), generó 8 falsos negativos, acertó en 60 de los 84 casos negativos (Verdaderos Negativos) y produjo 24 falsas alarmas. Traducido en métricas globales: Accuracy del 71,93%, F1-Score de 0,579, PR-AUC de 0,660 y una Tasa de Falsos Negativos del 26,67%. El análisis de importancia de características confirmó que la variable con mayor peso en la predicción fue ADL_x_FuncAss (importancia = 0,1179), seguida de MemoryComplaints (0,0984) y FunctionalAssessment (0,0936), lo cual es coherente con las correlaciones identificadas en la fase exploratoria y valida el enfoque de ingeniería de características adoptado. El análisis SHAP reforzó estos resultados al mostrar que valores altos de ADL_x_FuncAss desplazan consistentemente las predicciones hacia la clase positiva, confirmando que la brecha entre capacidad funcional y autonomía es el principal detonante del riesgo estimado.

Desde la perspectiva del tercer objetivo, la evaluación de usabilidad mediante demostración práctica y formulario estructurado permitió recoger percepción directa de los usuarios sobre la aplicación. Los resultados de esta evaluación complementan los indicadores técnicos del modelo, dado que una herramienta con métricas de clasificación aceptables pero baja usabilidad tendría un impacto real mínimo. En relación con la hipótesis del proyecto sobre que un registro digital organizado reduce la carga mental del cuidador y aumenta su sensación de seguridad, los resultados apuntan a una validación parcial: la aplicación AVI logra centralizar la información,

estructurar el registro conductual y presentar los resultados del modelo de forma comprensible para usuarios no técnicos, lo cual reduce la necesidad de mantener registros dispersos o depender exclusivamente de la memoria del cuidador. Sin embargo, la validación completa de la hipótesis requeriría un estudio longitudinal con un grupo mayor de cuidadores y pacientes reales.

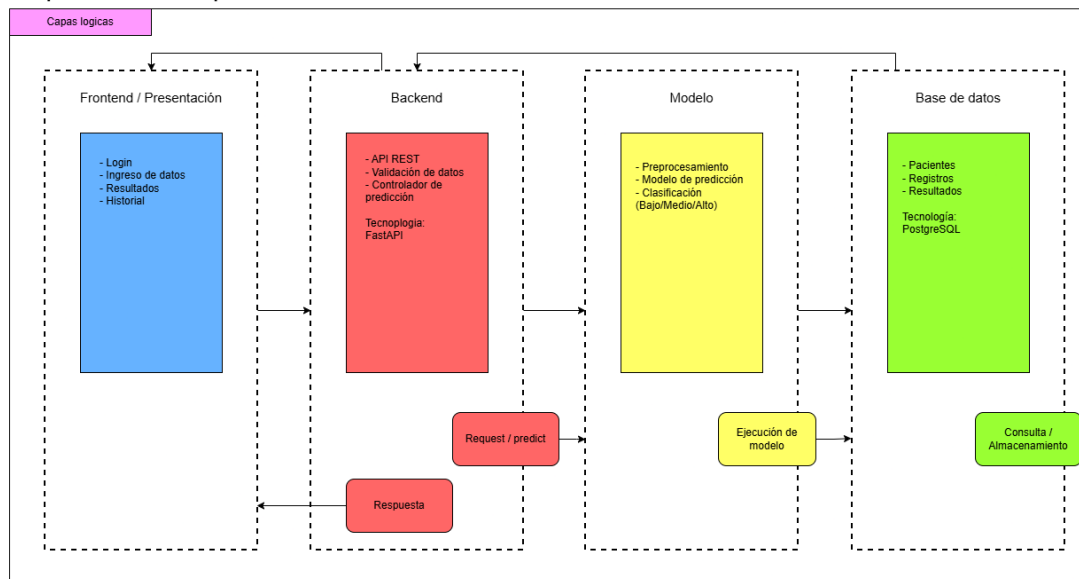
Propuesta de solución

AVI se define como un SaaS (Software as a Service) que tiene a los cuidadores, profesionales o no profesionales, de pacientes con Alzheimer como usuarios objetivo. La aplicación tiene tres funciones principales; una de ellas se basa en un modelo de machine learning que estima el porcentaje de riesgo de que un paciente de Alzheimer sufra una crisis de agitación. Los datos ingresados en el modelo y el resultado de este se registran en una base de datos para el seguimiento clínico de las estimaciones mediante un historial. La arquitectura de la aplicación está dividida en cuatro capas lógicas. La primera capa es el frontend, desarrollado en Python y Node.js que nos permite diseñar la parte visual de la aplicación , contando con alrededor de 13 pantallas, la segunda capa es el backend, basado en FastAPI para la gestión de peticiones REST que enviarán la información al modelo y a la base de datos, además de obtener la información médica, personal y registrada en el modelo para su visualización, la tercera capa es el modelo de machine learning y la última capa es la base de datos PostgreSQL alojada en la plataforma Neon, para el almacenamiento seguro del historial clínico y registros de estimación.

Figura 19

Arquitectura de la aplicación

Arquitectura de aplicación

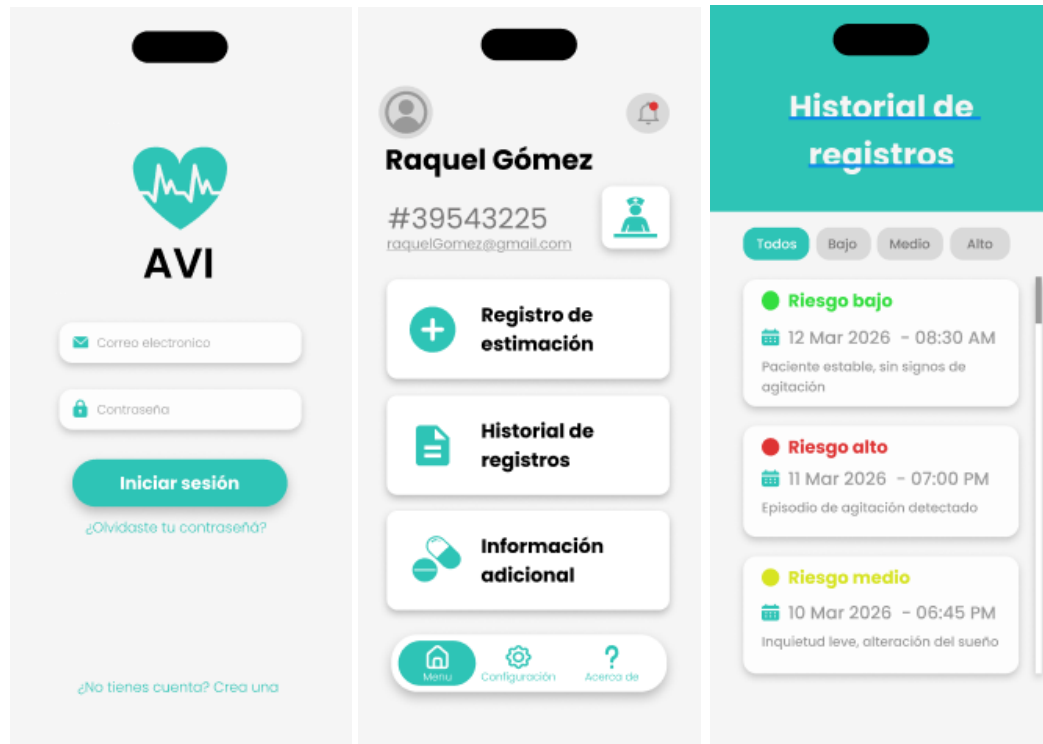


Nota: Imagen de elaboración propia desarrollada a partir de la plataforma draw.io

El diseño sigue principios de simplicidad e inclusión, utilizando interfaces de alto contraste y navegación intuitiva bajo el estándar WCAG 2.1. Esto busca facilitar el uso por parte de cuidadores, reduciendo la barrera tecnológica.

Figura 20

Interfaces principales de la aplicación AVI

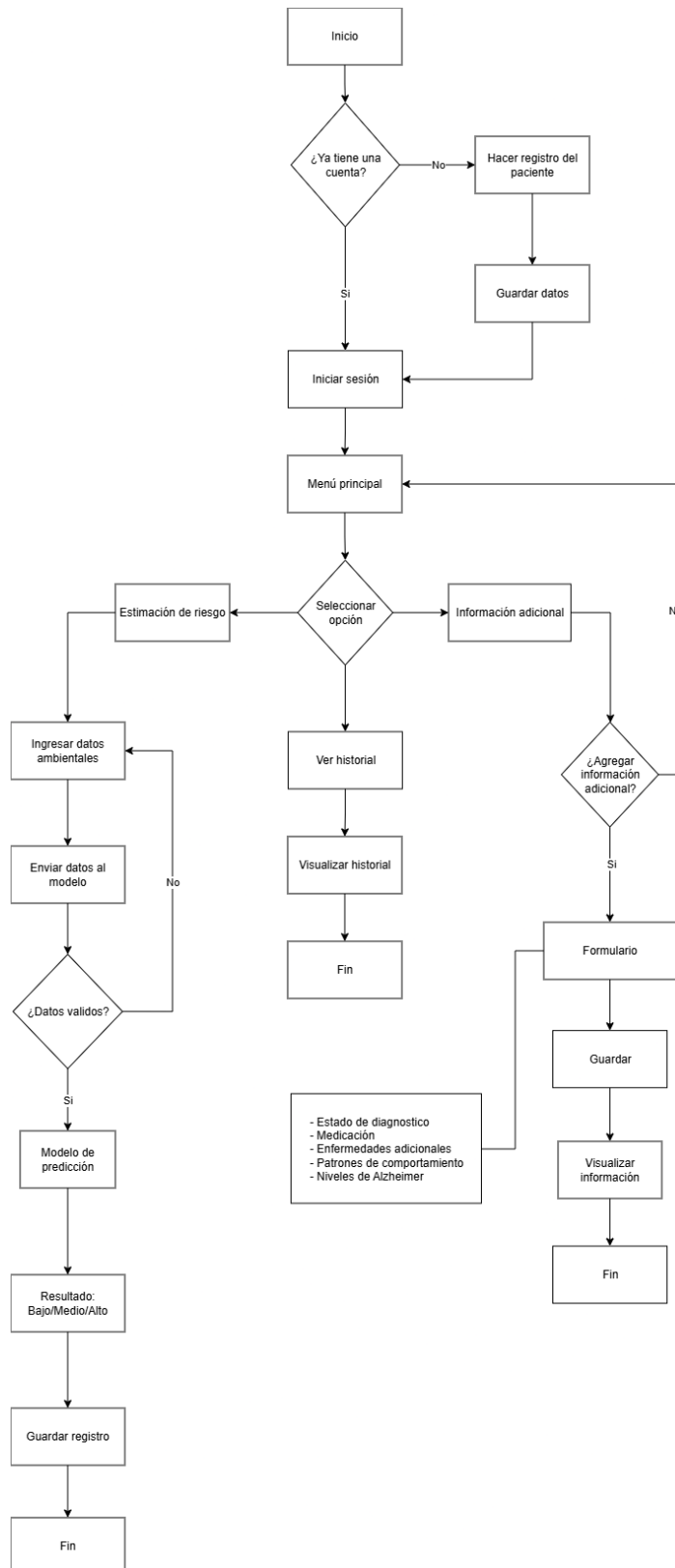


Nota: Imagen de elaboración propia desarrollada a partir de la plataforma Figma

El sistema permite registrar datos básicos del cuidador y del paciente, incluyendo niveles de Alzheimer y medicamentos. El núcleo de la solución es el módulo de estimación, que transforma datos conductuales en alertas de riesgo (alto, medio, bajo) mediante el modelo ya antes mencionado; los otros dos módulos u opciones son la visualización de registros e información adicional (información médica). Para el CRUD de la aplicación se creó una opción de configuración donde se pueden editar los pacientes y el propio usuario. Por último, un acerca de, que explica partes importantes del funcionamiento de la aplicación, como las tres calificaciones de los resultados del modelo o el funcionamiento de la aplicación en pocos pasos.

Figura 21

Flujo de la aplicación



Nota: Imagen de elaboración propia desarrollada a partir de la plataforma draw.io

Discusión

El desarrollo de AVI como herramienta tecnológica para la estimación anticipada de episodios de agitación en pacientes con Alzheimer permite situar los hallazgos de este proyecto en un contexto más amplio, tanto desde la perspectiva de la literatura científica revisada como desde las implicaciones prácticas del sistema construido.

El primer hallazgo de relevancia es que las variables funcionales y en particular la interacción entre las actividades de vida diaria y la evaluación funcional global (ADL_x_FuncAss) resultaron ser los predictores más importantes del riesgo de agitación, por encima de variables conductuales directas como Confusión o Desorientación. Esto está en línea con lo reportado por Fletcher-Lloyd et al. (2021), quienes identificaron que las actividades diarias capturadas mediante monitoreo en el hogar permiten anticipar el riesgo de agitación antes de que el episodio se manifieste. De forma similar, Ramesh et al. (2025) encontraron que variables de rutina como el sueño y las señales fisiológicas tienen capacidad predictiva sobre la agitación, lo cual refuerza la idea de que el deterioro en la autonomía cotidiana del paciente es una señal de alerta temprana más robusta que los síntomas conductuales explícitos. Estos datos evidencian que el modelo construido en este proyecto captura patrones que son consistentes con la evidencia internacional, a pesar de haber sido desarrollado con un dataset de escala reducida.

Un segundo hallazgo que merece análisis es la decisión de utilizar BehavioralProblems como variable objetivo en lugar de construir una definición compuesta de agitación. Esta elección, aunque pragmáticamente justificada dado que la alternativa reducía el ROC-AUC a $\sim 0,57$, introduce una circularidad metodológica que debe reconocerse con claridad. En los estudios revisados, como Bafaloukou et al. (2025) o Iaboni et al. (2022), la variable objetivo de agitación fue construida a partir de evaluaciones prospectivas y observaciones directas por parte de personal clínico, con separación temporal respecto a las variables predictoras. En el presente proyecto, ambas provienen de la misma instancia de observación, lo que limita la capacidad del modelo de representar genuinamente la predicción anticipada. A partir de ello se constata que esta es la

limitación metodológica más significativa del estudio, y que superarla requeriría una fase de recolección de datos longitudinales con pacientes reales en Colombia, donde el registro conductual de AVI podría convertirse precisamente en la fuente de datos para entrenar versiones futuras del modelo.

En tercer lugar, los valores de ROC-AUC obtenidos (0,747 en prueba) son comparables con los reportados en estudios similares que trabajan con datos clínicos estructurados. Mar et al. (2020) reportaron ROC-AUC entre 0,70 y 0,80 al predecir síntomas neuropsiquiátricos en pacientes con demencia usando Random Forest sobre datos de registros médicos reales. Gill et al. (2020) obtuvieron resultados similares combinando datos clínicos y de neuroimagen. Estos datos evidencian que el desempeño del modelo de AVI, aunque moderado, es consistente con lo que la literatura reporta para problemas de clasificación similares con tamaños de muestra comparables, y no constituye un resultado deficiente en términos absolutos, sino uno apropiado dado el alcance del proyecto.

Otro aspecto relevante es la elección del F2-score como criterio de optimización del umbral de decisión. Esta elección responde a una consideración clínica explícita: en el contexto del cuidado de pacientes con Alzheimer, un falso negativo, es decir, no alertar sobre un episodio que sí ocurrirá, tiene consecuencias más severas para el bienestar del paciente y el cuidador que una falsa alarma. Esta orientación hacia la sensibilidad está respaldada por Bafaloukou et al. (2025), quienes también priorizaron el Recall en su herramienta interpretable de detección de agitación en el hogar, argumentando que en entornos de cuidado informal la subdetección es el error más costoso. La decisión de mover el umbral de 0,50 a 0,28 sacrificó algo de precisión (de 0,640 a 0,478) pero ganó 20 puntos porcentuales en Recall (de 0,533 a 0,733), un intercambio que resulta favorable para el propósito de la herramienta.

Desde la perspectiva del desarrollo tecnológico, la implementación de AVI como MVP funcional en React Native con backend en FastAPI y base de datos en PostgreSQL demostró la viabilidad técnica de integrar un modelo de machine learning en un entorno móvil multiplataforma.

Esta arquitectura, aunque relativamente convencional, cumple con los principios de modularidad y escalabilidad necesarios para una eventual evolución del sistema. A diferencia de proyectos como AngelSense, que depende de hardware especializado, o CogniFit, orientado principalmente a la estimulación cognitiva, AVI se posiciona como una herramienta de registro y estimación de riesgo basada en variables conductuales que el cuidador ya observa en la cotidianidad, sin requerir dispositivos adicionales. Esta característica reduce la barrera de adopción y es coherente con lo que la literatura identifica como una necesidad no cubierta: herramientas sencillas, interpretables y no invasivas que puedan ser usadas por cuidadores no profesionales en el hogar (Bafaloukou et al., 2025; Fletcher-Lloyd et al., 2021).

Finalmente, en cuanto a la hipótesis del proyecto, los resultados obtenidos permiten sostener una validación parcial. La hipótesis planteaba que un registro digital organizado disminuye la carga mental del cuidador y aumenta su sensación de seguridad, independientemente de la complejidad técnica interna. La evaluación de usabilidad mostró que los usuarios percibieron la aplicación como comprensible y útil, lo que sugiere que la interfaz logra hacer accesible el modelo de estimación sin exponer al usuario a su complejidad técnica. No obstante, la comprobación robusta de la hipótesis en términos de reducción medible de la carga del cuidador requeriría un estudio controlado con seguimiento longitudinal, lo que excede el alcance de este proyecto, pero constituye una dirección natural de trabajo futuro.

Plan de divulgación

El plan de divulgación de AVI busca posicionar la herramienta tanto en el sector familiar como en el institucional. Para ello, se contempla una estrategia de marketing digital y presencial apoyada en campañas educativas, posicionamiento web y alianzas con asociaciones relacionadas con el Alzheimer y centros de salud. Asimismo, se proyecta la captación de usuarios mediante planes dirigidos a familias, clínicas y hospitales, garantizando la sostenibilidad financiera del proyecto. Para esto, AVI va a tener un plan de suscripción dividido en tres tipos de clientes: el primero, individual o familiar, con un costo de \$48.000 COP; segundo, para hogares geriátricos o

instituciones pequeñas, con un valor de \$800.000 COP; y para instituciones grandes como hospitales especializados, con un valor de \$4.000.000 COP.

Se va a plantear el establecimiento de la aplicación en Bogotá para el primer año, para el segundo y tercer año se va a hacer una expansión por toda Colombia. La inversión va a buscarse de manera mixta, una parte privada con una inversión de \$5.000.000 COP y una parte pública buscando fondos en iniciativas del gobierno o con créditos bancarios que abarcarían el 93% de la financiación. La tasa de retorno es de un 38,42% que nos permite recuperar la inversión en 3,34 años.

Figura 22

Pantalla de simulador financiero

INGRESOS/VENTAS DEL PRIMER AÑO					CRECIMIENTO PORCENTUAL EN VENTAS (CANTIDADES)				ANO BASE
NOMBRE DEL PRODUCTO O SERVICIO	CANTIDADES	PRECIO DE VENTA UNITARIO SIN IVA	INGRESOS TOTALES	ANO	2027	2028	2029	2030	2026
1 Suscripción Premium (Individual/Familiar)	1.200,00	\$ 48.000,00	\$ 57.600.000	43%	10,0%	15,0%	25,0%	30,0%	
2 Plan Clínica Pequeña (B2B)	36,00	\$ 800.000,00	\$ 28.800.000	21%	10,0%	15,0%	25,0%	30,0%	
3 Plan Institucional	12,00	\$ 4.000.000,00	\$ 48.000.000	36%	10,0%	15,0%	25,0%	30,0%	
4 Ventas de datos anonimizados	-	\$ -	\$ -	0%	0%	0%	0%	0%	
5	-	\$ -	\$ -	0%	0%	0%	0%	0%	
6	-	\$ -	\$ -	0%	0%	0%	0%	0%	
7	-	\$ -	\$ -	0%	0%	0%	0%	0%	
8	-	\$ -	\$ -	0%	0%	0%	0%	0%	
9	-	\$ -	\$ -	0%	0%	0%	0%	0%	
10	-	\$ -	\$ -	0%	0%	0%	0%	0%	
TOTAL			\$ 134.400.000	100%					

COSTOS DE CADA PRODUCTO O SERVICIO					PROYECCIONES			
NOMBRE DEL PRODUCTO O SERVICIO	CANTIDADES	COSTO UNITARIO DEL PRODUCTO O SERVICIO	COSTOS TOTALES	ANO	2027	2028	2029	2030
1 Suscripción Premium (Individual/Familiar)	1200	\$ 8.000,00	\$ 9.600.000	36%				
2 Plan Clínica Pequeña (B2B)	36	\$ 150.000,00	\$ 5.400.000	20%				
3 Plan Institucional	12	\$ 1.000.000,00	\$ 12.000.000	44%				
4 Ventas de datos anonimizados	0	\$ -	\$ -	0%				
5	0	\$ -	\$ -	0%				
6	0	\$ -	\$ -	0%				
7	0	\$ -	\$ -	0%				
8	0	\$ -	\$ -	0%				
9	0	\$ -	\$ -	0%				
10	0	\$ -	\$ -	0%				
TOTAL			\$ 27.000.000	100%				

ANO	2026	2027	2028	2029	2030
VENTAS ANUALES	\$ 134.400.000,0	\$ 153.753.600,0	\$ 184.773.388,8	\$ 238.357.671,6	\$ 321.639.842,0
COSTOS ANUALES	\$ 27.000.000,0	\$ 29.700.000,0	\$ 34.155.000,0	\$ 42.693.750,0	\$ 55.501.875,0
MARGEN OPERATIVO	\$ 107.400.000,0	\$ 124.053.600,0	\$ 150.618.388,8	\$ 195.663.921,6	\$ 266.137.967,0

Nota: Imagen de elaboración propia de la plantilla de simulador financiero en excel

Complementariamente, se plantea una estrategia de divulgación científica y social basada en tutoriales interactivos, accesibilidad bajo estándares WCAG 2.1 y la presentación de resultados académicos y técnicos del proyecto, con el fin de fomentar la adopción de la herramienta y contribuir a la mejora de la calidad de vida de pacientes, cuidadores y profesionales de la salud.

Conclusiones

El desarrollo de este proyecto permitió demostrar que es posible construir una herramienta tecnológica funcional para la estimación anticipada de episodios de agitación en pacientes con Alzheimer, integrando técnicas de aprendizaje automático con una aplicación móvil orientada al cuidador. Los resultados obtenidos en cada uno de los objetivos específicos permiten formular las siguientes conclusiones.

En relación con el primer objetivo, el análisis cuantitativo del dataset clínico dentro del subgrupo de pacientes con Alzheimer confirmado permitió identificar un conjunto de variables conductuales, funcionales y cognitivas con asociación real y medible con los episodios de agitación. Las actividades de vida diaria (ADL), la evaluación funcional global, las quejas de memoria y las variables de confusión y desorientación se consolidaron como los predictores más relevantes, con correlaciones entre 5 y 12 veces más fuertes que en el dataset completo. Este hallazgo confirma que el deterioro funcional progresivo, más que los síntomas conductuales aislados, es la señal de alerta más robusta disponible en registros clínicos estructurados, y que su captura sistemática mediante un instrumento digital como AVI tiene sustento metodológico sólido.

En cuanto al segundo objetivo, el modelo Gradient Boosting con calibración sobre datos originales y umbral optimizado mediante F2-score alcanzó un ROC-AUC de 0,747 y un Recall del 73,33% sobre el conjunto de prueba, con una tasa de falsos negativos del 26,67%. Estos resultados, aunque moderados, son consistentes con los reportados en la literatura internacional para problemas de clasificación similares con tamaños de muestra comparables, y son adecuados para el propósito de apoyo al seguimiento que persigue la aplicación. El proceso de implementación demostró, además, que decisiones metodológicas como el manejo correcto de la calibración y la elección de la métrica de optimización del umbral tienen un impacto determinante sobre el desempeño real del modelo, más allá de la elección del algoritmo en sí.

Respecto al tercer objetivo, la evaluación del desempeño del modelo y la usabilidad de la aplicación confirmaron que AVI es un prototipo funcional viable. La comunicación entre frontend, backend, modelo y base de datos operó de forma correcta en todas las pruebas realizadas, y los

usuarios que participaron en la evaluación de usabilidad percibieron la herramienta como comprensible y útil para el seguimiento de sus pacientes. Esto respalda la hipótesis del proyecto en su dimensión de accesibilidad: la complejidad técnica interna del modelo no representa una barrera para el cuidador, dado que la interfaz transforma los datos en alertas de riesgo interpretables sin requerir conocimiento técnico.

El proyecto tiene dos limitaciones principales que deben reconocerse. La primera es el tamaño reducido del dataset de entrenamiento, con apenas 147 casos positivos en el conjunto de entrenamiento, lo que limita la capacidad de generalización del modelo. La segunda es la circularidad en la construcción de la variable objetivo, que al derivarse del mismo dataset que las características predictoras, introduce una dependencia metodológica que no replica fielmente el escenario de predicción prospectiva. Estas limitaciones no invalidan los resultados, pero sí definen con claridad el alcance del prototipo como herramienta de screening y no como sistema de diagnóstico clínico definitivo.

Como oportunidad de trabajo futuro, la aplicación AVI constituye en sí misma el instrumento más valioso para superar estas limitaciones: a medida que los cuidadores registren datos reales de manera longitudinal, será posible construir un dataset propio, con separación temporal entre variables predictoras y etiquetas de agitación, sobre el cual entrenar versiones mejoradas del modelo. Esto abriría la posibilidad de desarrollar un sistema verdaderamente predictivo, adaptado al contexto colombiano y ajustado a las particularidades de los pacientes y cuidadores locales, contribuyendo así a cerrar la brecha entre la investigación experimental en este campo y las herramientas prácticas disponibles para quienes cuidan a diario a personas con Alzheimer.

Referencias bibliográficas

Bafaloukou, M., Schalkamp, A.-K., Fletcher-Lloyd, N., Capstick, A., Walsh, C., Sandor, C., Kouchaki, S., Group, C., Nilforooshan, R., & Barnaghi, P. (2024). An Interpretable Machine Learning Tool for In-Home Screening of Agitation Episodes in People Living

with Dementia. *MedRxiv*, 2024.07.30.24311178.

<https://doi.org/10.1101/2024.07.30.24311178>

Fletcher-Lloyd, N., Soreq, E., Wilson, D., Nilforooshan, R., Sharp, D. J., & Barnaghi, P. (2021).

Home monitoring of daily living activities and prediction of agitation risk in a cohort of people living with dementia. *Alzheimer's & Dementia*, 17(S12), e058614.

<https://doi.org/https://doi.org/10.1002/alz.058614>

HekmatiAthar, S., Goins, H., Samuel, R., Byfield, G., & Anwar, M. (2021). Data-Driven

Forecasting of Agitation for Persons with Dementia: A Deep Learning-Based Approach.

SN Computer Science, 2(4), 326. <https://doi.org/10.1007/s42979-021-00708-3>

Homdee, N., Alam, R., Hayes, J. A., Hamid, T., Park, J., Wolfe, S., Goins, H., Fyffe, N., Newbold,

T., Smith-Jackson, T., Bankole, A., Anderson, M. S., & Lach, J. (2019). Agitation

Monitoring and Prevention System for Dementia Caregiver Empowerment. *Computer*,

52(11), 30–39. <https://doi.org/10.1109/MC.2019.2933192>

Oliveira, R., Feres, R., Barreto, F., & Abreu, R. (2022). CNN for Elderly Wandering Prediction in

Indoor Scenarios. *SN Computer Science*, 3(3), 230. [https://doi.org/10.1007/s42979-022-](https://doi.org/10.1007/s42979-022-01091-3)

[01091-3](https://doi.org/10.1007/s42979-022-01091-3)

Ramesh, K., Yakoub, A., Ghoneim, Y., Al Korabi, R., Ramesh, J., Sagahyroon, A., & Aloul, F.

(2025). Machine Learning Prediction of Agitation in Dementia Patients Using Sleep and

Physiological Data. *Applied Sciences*, 15(18). <https://doi.org/10.3390/app15189908>

Zhang, K., Yuan, H., & Fang, Y. (2022). Evaluation of Smart Agitation Prediction and

Management for Dementia Care and Novel Universal Village Oriented Solution for

Integration, Resilience, Inclusiveness and Sustainability. *2022 6th International*

Conference on Universal Village (UV), 1–34.

<https://doi.org/10.1109/UV56588.2022.10185497>

Bafaloukou, M., Schalkamp, A.-K., Fletcher-Lloyd, N., Capstick, A., Walsh, C., Sandor, C.,

Kouchaki, S., Group, C., Nilforooshan, R., & Barnaghi, P. (2024). An Interpretable

Machine Learning Tool for In-Home Screening of Agitation Episodes in People Living with Dementia. *MedRxiv*, 2024.07.30.24311178.

<https://doi.org/10.1101/2024.07.30.24311178>

Algase, D. L., Beck, C., Kolanowski, A., Whall, A., Berent, S., Richards, K., & Beattie, E. (1996).

Need-driven dementia-compromised behavior: An alternative view of disruptive behavior. *American Journal of Alzheimer's Disease*, 11(6), 10-

19. <https://doi.org/10.1177/153331759601100603>

Gómez, L. (11 de septiembre de 2024). Francisco Lopera: Un faro para el Alzheimer y las ciencias

de la salud. *Pesquisa Javeriana*. <https://www.javeriana.edu.co/pesquisa/francisco-lopera-faro-otra-ciencia/>

Konexión Alzheimer. (2026). *Innovación tecnológica en el cuidado del Alzheimer: herramientas*

que mejoran la calidad de vida. <https://konexionalzheimer.com/innovacion-tecnologica-en-el-cuidado-del-alzheimer-herramientas-que-mejoran-la-calidad-de-vida/>

Mauleon, A., Ismail, Z., Rosenberg, P. B., Tariot, P. N., Lyketsos, C. G., & Gauthier, S. (2021).

Agitation in Alzheimer's disease: Novel outcome measures reflecting the International Psychogeriatric Association (IPA) agitation criteria. *Alzheimer's & Dementia*,

17(S10). <https://doi.org/10.1002/ALZ.12335>

Meunier, J., Creel, K., Loubert, A., & de Mauleon, A. (2024). Defining a clinically meaningful

within-patient change threshold for the Cohen-Mansfield Agitation Inventory in Alzheimer's dementia. *Frontiers in Neurology*,

15. <https://doi.org/10.3389/fneur.2024.1379062>

Ministerio de Salud y Protección Social. (21 de septiembre de 2020). *Colombia cuenta con ley que*

garantiza la atención del Alzheimer. <https://www.minsalud.gov.co/Paginas/Colombia-cuenta-con-ley-que-garantiza-la-atencion-del-Alzheimer.aspx>

Organización Mundial de la Salud. (31 de marzo de

2025). *Demencia*. <https://www.who.int/es/news-room/fact-sheets/detail/dementia>

Addae, S., Kim, J., Smith, A., Rajana, P., & Kang, M. (2024). Smart solutions for detecting, predicting, monitoring, and managing dementia in the elderly: A survey. *IEEE Access*, 12, 100026–100056. <https://doi.org/10.1109/ACCESS.2024.3421966>

Au-Yeung, W. T. M., Miller, L., Beattie, Z., Dodge, H. H., Reynolds, C., Vahia, I., & Kaye, J. (2020). Sensing a problem: Proof of concept for characterizing and predicting agitation. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 6(1), e12079.

Bafaloukou, M., Schalkamp, A.-K., Fletcher-Lloyd, N., Capstick, A., Walsh, C., Sandor, C., Kouchaki, S., Nilforooshan, R., & Barnaghi, P. (2025). An interpretable machine learning tool for in-home monitoring of agitation episodes in people living with dementia: A proof-of-concept study. *eClinicalMedicine*, 80, 103032. <https://doi.org/10.1016/j.eclinm.2024.103032>

Bijlani, N., Maldonado, O. M., Nilforooshan, R., Kouchaki, S., & Barnaghi, P. (2024). Utilizing graph neural networks for adverse health detection and personalized decision making in sensor-based remote monitoring for dementia care. *Computers in Biology and Medicine*, 183, 109287.

Cheung, J. C.-W., So, B. P.-H., Ho, K. H. M., Wong, D. W.-C., Lam, A. H.-F., & Cheung, D. S. K. (2022). Wrist accelerometry for monitoring dementia agitation behaviour in clinical settings: A scoping review. *Frontiers in Psychiatry*, 13, 913213. <https://doi.org/10.3389/fpsyt.2022.913213>

Delmastro, F., Di Martino, F., & Dolciotti, C. (2020). Cognitive training and stress detection in mild cognitive impairment frail older people through wearable sensors and machine learning. *IEEE Access*, 8, 65573–65584. <https://doi.org/10.1109/ACCESS.2020.2985301>

Fletcher-Lloyd, N., Soreq, E., Wilson, D., Nilforooshan, R., Sharp, D. J., & Barnaghi, P. (2021). Home monitoring of daily living activities and prediction of agitation risk in a cohort of people living with dementia. *Alzheimer's & Dementia*, 17(S12), e058614. <https://doi.org/10.1002/alz.058614>

Fernandes, S., Rosselet Amoussou, J., Gomes da Rocha, C., Perruchoud, E., von Gunten, A., Mabire, C., & Verloo, H. (2025). Using artificial intelligence–based technologies for the early detection of behavioral and psychological symptoms of dementia: Scoping review. *JMIR Aging*, 8, e76074. <https://doi.org/10.2196/76074>

Favela, J., Cruz-Sandoval, D., Morales-Tellez, A., & Lopez-Nava, I. H. (2020). Monitoring behavioral symptoms of dementia using activity trackers. *Journal of Biomedical Informatics*, 109, 103520.

Goerss, D., Hein, A., Bader, S., Halek, M., Kernebeck, S., Kutschke, A., Heine, C., Krueger, F., Kirste, T., & Teipel, S. (2020). Automated sensor-based detection of challenging behaviors in advanced stages of dementia in nursing homes. *Alzheimer's & Dementia*, 16(5), 672–680.

Hasan, Z., Shum, L. C., Faruk, T., McGilton, K. S., Chu, C. H., McArthur, C., Flint, A. J., Mihailidis, A., Khan, S. S., & Iaboni, A. (2025). Machine learning to detect motor agitation in people with dementia using real-time location data. *American Journal of Geriatric Psychiatry*, 33(11), 1171–1183. <https://doi.org/10.1016/j.jagp.2025.06.001>

Iaboni, A., Spasojevic, S., Newman, K., Schindel Martin, L., Wang, A., Ye, B., Mihailidis, A., & Khan, S. S. (2022). Wearable multimodal sensors for the detection of behavioral and psychological symptoms of dementia using personalized machine learning models. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 14(1), e12305.

Kim, J., Cheon, S., & Lim, J. (2022). IoT-based unobtrusive physical activity monitoring system for predicting dementia. *IEEE Access*, 10, 26078–26089. <https://doi.org/10.1109/ACCESS.2022.3156607>

Kim, K., Lee, S., Kim, S., Kim, J., Shin, D., & Shin, D. (2020). Sensor-based deviant behavior detection system using deep learning to help dementia caregivers. *IEEE Access*, 8, 136004–136015. <https://doi.org/10.1109/ACCESS.2020.3011654>

Khan, S. S., Mishra, P. K., Javed, N., Ye, B., Newman, K., Mihailidis, A., & Iaboni, A. (2022). Unsupervised deep learning to detect agitation from videos in people with dementia. *IEEE Access*, 10, 10349–10360. <https://doi.org/10.1109/ACCESS.2022.3143990>

Khan, S. S., Mishra, P. K., Ye, B., Patel, S., Newman, K., Mihailidis, A., & Iaboni, A. (2025). A novel multi-modal sensor dataset and benchmark to detect agitation in people living with dementia. *ACM Transactions on Computing for Healthcare*.

Liu, X., Chau, K. Y., Zheng, J., Deng, D., & Tang, Y. M. (2024). Artificial intelligence approach for detecting and classifying abnormal behaviour in older adults using wearable sensors. *Journal of Rehabilitation and Assistive Technologies Engineering*, 11, 1–18. <https://doi.org/10.1177/20556683241288459>

Malioukis, A., Snead, R. S., Marczika, J., Ambalavanan, R., Towett, G., & Mbogori-Kairichi, M. (2026). Digital biomarkers for early agitation detection in dementia. *Frontiers in Neurology*, 17, 1683517.

Mar, J., Gorostiza, A., Ibarrondo, O., Cernuda, C., Arrospide, A., Iruin, A., Larranaga, I., Tainta, M., Ezpeleta, E., & Alberdi, A. (2020). Validation of random forest machine learning models to predict dementia-related neuropsychiatric symptoms in real-world data. *Journal of Alzheimer's Disease*, 77(3), 855–864. <https://doi.org/10.3233/JAD-200345>

Ranson, J. M., Bucholc, M., Lyall, D., Newby, D., Winchester, L., Oxtoby, N. P., Veldsman, M., Rittman, T., Marzi, S., Skene, N., Al Khleifat, A., Foote, I. F., Orgeta, V., Kormilitzin, A., Lourida, I., & Llewellyn, D. J. (2023). Harnessing the potential of machine learning and artificial

intelligence for dementia research. *Brain Informatics*, 10, 6. <https://doi.org/10.1186/s40708-022-00183-3>

Rezvani, R., Kouchaki, S., Nilforooshan, R., Sharp, D. J., & Barnaghi, P. (2021). Analysing behavioural changes in people with dementia using in-home monitoring technologies. *Alzheimer's & Dementia*, 17(S1).

Rykov, Y. G., Ng, K. P., Patterson, M. D., & Gangwar, B. A. (2024). Predicting the severity of mood and neuropsychiatric symptoms from digital biomarkers using wearable physiological data and deep learning. *Computers in Biology and Medicine*, 180, 108959.

Xie, B., Tao, C., Li, J., Hilsabeck, R. C., & Aguirre, A. (2020). Artificial intelligence for caregivers of persons with Alzheimer's disease and related dementias: Systematic literature review. *JMIR Medical Informatics*, 8(8), e18189. <https://doi.org/10.2196/18189>

Yuan, H., Yang, T., Xie, Q., Lledos, G., Chou, W.-H., & Yu, W. (2024). Modeling and mobile home monitoring of behavioral and psychological symptoms of dementia (BPSD). *BMC Psychiatry*, 24, 197. <https://doi.org/10.1186/s12888-024-05579-5>

World Health Organization. (2023). Dementia. <https://www.who.int/news-room/fact-sheets/detail/dementia>