



TRABAJO FIN DE GRADO — BUSINESS ANALYTICS

Curso 2025-2026

PROYECTO KORA: DISEÑO Y DESARROLLO DE APIs PARA LA MONITORIZACIÓN DE MAYORES

Autor:

Enrique Quintero Sieira

Director:

José Portela González

Resumen

El envejecimiento de la población española plantea un reto creciente: cada vez más personas mayores viven solas y sus familias carecen de información sobre cómo se encuentran en el día a día. Las soluciones actuales de teleasistencia son reactivas, ya que actúan cuando el problema ya ha ocurrido. Kora propone un enfoque anticipatorio basado en el análisis de datos biométricos para detectar desviaciones del patrón habitual antes de que se conviertan en una urgencia.

Este Trabajo Fin de Grado se centra en la pieza analítica que sostiene esa promesa: un sistema de detección de anomalías sobre señales biométricas. El trabajo desarrolla el modelo de detección con la técnica Isolation Forest, lo valida sobre un conjunto de datos simulado que reproduce patrones fisiológicos realistas, y lo expone a través de una interfaz de programación sencilla que permite recibir medidas y devolver alertas clasificadas por nivel de gravedad. La gestión del envío de avisos se resuelve con una herramienta de automatización no-code. El trabajo se completa con un análisis del sector, un encuadre del modelo de negocio y una evaluación de la viabilidad técnica y económica de la propuesta.

El proyecto se ha desarrollado en el marco del concurso Comillas Emprende durante el curso 2025-2026. El resultado demuestra que un sistema de monitorización predictiva con calidad suficiente para una primera fase de validación puede construirse con recursos limitados y tecnologías abiertas.

Palabras clave: Kora, detección de anomalías, Isolation Forest, wearables, eldercare, análisis de datos, healthtech, Silver Economy, no-code.

Abstract

The ageing of the Spanish population poses a growing challenge: more and more elderly people live alone, and their families lack information about how they are doing on a daily basis. Current telecare solutions are reactive, acting only once a problem has already occurred. Kora proposes an anticipatory approach based on the analysis of biometric data to detect deviations from the usual pattern before they escalate into an emergency.

This dissertation focuses on the analytical core that supports that promise: an anomaly detection system applied to biometric signals. The work develops the detection model using the Isolation Forest technique, validates it on a simulated dataset that reproduces realistic physiological patterns, and exposes it through a simple programming interface that receives measurements and returns alerts classified by severity level. The dispatch of notifications is handled with a no-code automation tool. The work is completed with a sector analysis, a framing of the business model, and an assessment of the technical and economic feasibility of the proposal.

The project was developed within the Comillas Emprande entrepreneurship competition during the 2025-2026 academic year. The result shows that a predictive monitoring system of sufficient quality for an initial validation phase can be built with limited resources and open technologies.

Keywords: Kora, anomaly detection, Isolation Forest, wearables, eldercare, data analysis, healthtech, Silver Economy, no-code.

Índice

Resumen.....	2
Abstract.....	3
Índice.....	4
1. Introducción	7
1.1. Origen y justificación del proyecto	7
1.2. Objetivos	7
1.3. Pregunta de investigación	8
1.4. Metodología y plan de trabajo	8
1.5. Estructura del documento	9
2. Marco teórico y revisión de la literatura	11
2.1. Envejecimiento demográfico y Silver Economy	11
2.2. Wearables y monitorización remota de salud	12
2.3. Detección de anomalías en señales biométricas	12
2.4. Exposición de modelos mediante interfaces de programación	13
3. Análisis del entorno y oportunidad	15
3.1. Contextualización del problema	15
3.2. Análisis PESTEL	15
3.3. Análisis competitivo y posicionamiento	16
4. Modelo de negocio.....	18
4.1. Propuesta de valor.....	18
4.2. Segmentos de cliente.....	18
4.3. Fuentes de ingresos y estructura de costes.....	18
4.4. Encaje con el perfil del grado	19
5. Datos y metodología analítica.....	20
5.1. Origen y naturaleza de los datos	20
5.2. Generación del conjunto simulado.....	20
5.3. Análisis exploratorio	21
5.4. Construcción de la línea base individual	21
6. Modelo de detección de anomalías	23
6.1. Elección del algoritmo: Isolation Forest	23
6.2. Variables y entrenamiento	23
6.3. Clasificación por niveles de gravedad	25
6.4. Evaluación del modelo.....	25
7. Exposición del modelo: la API de monitorización	27
7.1. Propósito y alcance	27

7.2. Endpoints y flujo de datos.....	27
7.3. Decisiones de diseño.....	27
8. Gestión de alertas con herramientas no-code.....	29
8.1. Justificación del enfoque no-code.....	29
8.2. Diseño del flujo de notificaciones.....	29
8.3. Reglas de escalado	29
9. Evaluación de la solución	31
9.1. Resultados del modelo	31
9.2. Viabilidad económica	31
9.3. Limitaciones.....	32
10. Conclusiones	34
10.1. Cumplimiento de objetivos.....	34
10.2. Aportaciones	34
10.3. Líneas futuras.....	34
Bibliografía	36
Anexos	39

1. Introducción

1.1. Origen y justificación del proyecto

El interés por este proyecto parte de una observación sencilla: la dificultad de saber, sin invadir, cómo se encuentra realmente una persona mayor que vive sola. España envejece a un ritmo que las cifras documentan con precisión, pero el dato cuantitativo no refleja por sí solo la realidad cotidiana del cuidado a distancia. A esa realidad social se suma una circunstancia tecnológica favorable. Los dispositivos vestibles han alcanzado un grado de madurez que permite medir constantes vitales de forma continua y a un coste asequible, y las técnicas de detección de anomalías sobre datos han demostrado capacidad para identificar señales tempranas de deterioro fisiológico.

Sobre esa base se construye Kora, un sistema pensado para personas mayores que viven solas o con su pareja. La idea completa contempla tres componentes: un dispositivo vestible adaptado al segmento, una aplicación móvil dirigida al familiar y una capa de inteligencia que analiza los datos para anticipar problemas. De estos tres componentes, el presente trabajo aborda en profundidad el tercero, que es donde reside el valor diferencial del producto y donde las competencias del Grado en Análisis de Negocios encuentran su aplicación más directa.

El alcance de este Trabajo Fin de Grado abarca dos planos. El primero es analítico: el diseño, entrenamiento y evaluación de un modelo de detección de anomalías sobre señales biométricas. El segundo es de encuadre: un análisis del sector, una caracterización del modelo de negocio y una valoración de la viabilidad técnica y económica de la solución. El desarrollo se ha realizado en el marco del concurso Comillas Emprende durante el curso 2025-2026, siguiendo la metodología Elevatorfy.

La justificación del trabajo es doble. Por un lado, demostrar que la promesa de anticipación de Kora se sostiene sobre una base técnica real y construible con recursos modestos. Por otro, mostrar que el perfil de Análisis de Negocios resulta idóneo para abordar un problema de estas características, en el que el tratamiento de datos y la modelización se ponen al servicio de una decisión empresarial concreta.

1.2. Objetivos

El objetivo general consiste en diseñar, entrenar y evaluar un sistema de detección de anomalías sobre datos biométricos que permita identificar señales tempranas de deterioro en personas mayores, y exponerlo de forma que pueda integrarse en el ecosistema Kora.

De este objetivo general se derivan los siguientes objetivos específicos:

- **Analizar el sector** del cuidado de mayores y la Silver Economy en España, para dimensionar el problema y entender las limitaciones de las soluciones actuales.
- **Encuadrar el modelo de negocio** de Kora en su propuesta de valor, sus segmentos de cliente y su estructura de ingresos y costes.
- **Generar un conjunto de datos simulado** que reproduzca patrones fisiológicos realistas de una persona mayor, con anomalías controladas para validar el sistema.
- **Construir un modelo de detección de anomalías** basado en la técnica Isolation Forest, ajustado al patrón individual de cada usuario.
- **Definir una clasificación de las alertas** por niveles de gravedad que evite la fatiga por exceso de notificaciones.
- **Exponer el modelo** a través de una interfaz de programación sencilla que reciba medidas y devuelva el estado del usuario.
- **Resolver el envío de alertas** mediante una herramienta de automatización no-code que orqueste los avisos a los familiares.
- **Evaluar la solución** en sus dimensiones técnica (rendimiento del modelo) y económica (coste de operación).

1.3. Pregunta de investigación

La pregunta que vertebra el trabajo se formula así:

¿Es posible construir un sistema de detección de anomalías sobre señales biométricas de personas mayores, con una calidad suficiente para una primera fase de validación y un coste de operación compatible con un modelo de negocio sostenible, empleando técnicas analíticas estándar y herramientas abiertas?

La pregunta no tiene una respuesta evidente de antemano. La detección de patrones fisiológicos anómalos es un problema bien estudiado en la literatura clínica, pero su aplicación a un contexto de bajo coste, sobre el patrón individual de cada usuario y con interpretabilidad suficiente para generar confianza en las familias, plantea retos de diseño concretos. El trabajo aborda esos retos y valora hasta qué punto la respuesta es afirmativa.

1.4. Metodología y plan de trabajo

El trabajo se ha organizado en tres bloques sucesivos, cada uno con su enfoque propio.

Bloque 1. Análisis y encuadre. La primera fase comprende la revisión de literatura y de fuentes oficiales para dimensionar el problema del envejecimiento, estudiar las tendencias del sector y mapear el panorama competitivo. Las herramientas empleadas son las del análisis estratégico: PESTEL, fuerzas competitivas y Business Model Canvas, aplicadas con un criterio funcional orientado a fundamentar las decisiones de diseño posteriores.

Bloque 2. Desarrollo analítico. La segunda fase es el núcleo del trabajo. Comprende la generación del conjunto de datos simulado, el análisis exploratorio de las señales, la construcción de la línea base individual de cada usuario y el entrenamiento del modelo de detección de anomalías con la técnica Isolation Forest. Todo el desarrollo se ha realizado en Python con las librerías habituales de análisis de datos: pandas para la manipulación, scikit-learn para el modelado y matplotlib para la visualización.

Bloque 3. Exposición y evaluación. La tercera fase expone el modelo a través de una interfaz de programación sencilla, resuelve el envío de alertas con una herramienta no-code y evalúa el conjunto en sus dimensiones técnica y económica. La evaluación técnica mide la capacidad del modelo para detectar anomalías sobre el conjunto simulado. La económica estima el coste de operación a distintos niveles de uso.

El trabajo se apoya de forma deliberada en el instrumental que ofrece el Grado en Análisis de Negocios. El tratamiento de datos se realiza con la librería pandas, que permite manipular tablas de medidas con facilidad. El modelado emplea scikit-learn, la librería de referencia para aprendizaje automático en Python, ampliamente trabajada en las asignaturas de la titulación. La visualización se apoya en matplotlib. Esta elección de herramientas no es casual: prioriza la transparencia y la capacidad de explicar cada paso del proceso, frente a soluciones más opacas que dificultarían la defensa de los resultados.

El cronograma se ha ajustado al calendario de entregas de la Guía Docente: propuesta inicial en octubre de 2025, entregas intermedias en diciembre y febrero, y entrega final en abril de 2026. Los hitos del concurso Comillas Emprende han servido como puntos de control externos.

1.5. Estructura del documento

El documento se organiza en diez capítulos. Tras esta introducción, el segundo capítulo presenta el marco teórico. El tercero analiza el entorno y la oportunidad de negocio. El cuarto encuadra el modelo de negocio. El quinto describe los datos y la metodología analítica empleada. El sexto desarrolla el modelo de detección de anomalías, que constituye el centro del trabajo. El séptimo explica cómo se expone ese modelo a través de una interfaz de

programación. El octavo aborda la gestión de alertas con herramientas no-code. El noveno recoge la evaluación de la solución y el décimo cierra con las conclusiones y las líneas futuras. La bibliografía y los anexos completan el documento.

2. Marco teórico y revisión de la literatura

El trabajo se sitúa en el cruce de tres campos: la economía del envejecimiento, los wearables aplicados a la salud y la detección de anomalías sobre datos. Este capítulo revisa los conceptos y la literatura relevantes en cada uno, sin pretensión de exhaustividad, con el fin de dejar asentadas las bases sobre las que se construyen las decisiones de los capítulos posteriores.

2.1. Envejecimiento demográfico y Silver Economy

El envejecimiento de la población es una de las tendencias mejor documentadas de las últimas décadas. España figura, junto con Italia y Japón, entre los países donde el proceso avanza con mayor rapidez. Según el informe de Pérez Díaz y otros para el Consejo Superior de Investigaciones Científicas (2024), la población mayor de 65 años supera ya los 9,7 millones de personas, más del 20% del total, y las proyecciones apuntan a que superará el 30% hacia 2050.

Lo que transforma el problema no es solo el número, sino la estructura del cuidado. La convivencia intergeneracional, que hace décadas era la norma, hoy es la excepción. Más de 1,7 millones de personas mayores de 70 años viven solas en España. Cuando existe cuidado, recae mayoritariamente en hijos de entre 45 y 60 años que residen en otra ciudad y compaginan esa responsabilidad con su propia vida laboral y familiar. La Silver Economy, que la Comisión Europea cifró en torno a 5,7 billones de euros en 2025, se construye sobre esta tensión: hay demanda solvente de soluciones que alivien la carga del cuidado, y la oferta no termina de estar a la altura.

La literatura económica reciente coincide en que el modelo de cuidado institucional está saturado y el modelo de cuidado informal resulta insostenible al ritmo actual. La alternativa que se apunta con más fuerza es el envejecimiento en el domicilio asistido por tecnología. Ahora bien, la teleasistencia clásica española, basada en el botón de emergencia, cubre lo básico pero no responde a la pregunta más relevante para una familia, que no es qué hacer si ocurre algo grave, sino cómo saber cómo se encuentra su familiar cada día.

El concepto de Silver Economy abarca mucho más que el cuidado sanitario. Comprende el conjunto de actividades económicas vinculadas a las personas mayores: vivienda adaptada, turismo, ocio, servicios financieros, formación y tecnología asistencial, entre otras. Su relevancia no deja de crecer porque el segmento de población al que se dirige no solo aumenta en número, sino también en capacidad de gasto y en esperanza de vida con autonomía. En este marco, las soluciones tecnológicas de apoyo al envejecimiento en el domicilio ocupan un lugar

destacado, al situarse en la intersección entre la necesidad social, la sostenibilidad del sistema sanitario y la oportunidad de mercado.

2.2. Wearables y monitorización remota de salud

Conviene distinguir dos categorías de dispositivos vestibles que a menudo se confunden. Por un lado están los wearables generalistas, como Apple Watch, Garmin o Fitbit, que han incorporado funciones de salud sobre un producto concebido para deporte y actividad diaria. Por otro están los wearables de orientación clínica, diseñados desde el principio para uso médico y validados con población clínica.

La diferencia tiene consecuencias prácticas. El estudio Apple Heart, con más de 400.000 participantes (Perez y otros, 2019), mostró que el algoritmo de detección de fibrilación auricular del Apple Watch alcanza una sensibilidad y especificidad aceptables en población general, pero su comportamiento se degrada en perfiles atípicos, como personas mayores polimedizadas. Esto es justamente lo que Kora busca abordar: no fabricar un wearable mejor, sino construir una capa de inteligencia entrenada con el patrón de la población diana y, sobre todo, con el patrón individual de cada usuario.

La referencia internacional más citada en este terreno es CarePredict, una solución estadounidense que despliega dispositivos vestibles en residencias asistidas y que ha publicado reducciones de hospitalizaciones gracias a la detección temprana. Es una referencia útil porque demuestra que el enfoque funciona, aunque opera en residencias y no en domicilios, con un modelo distinto al planteado para Kora.

Más allá de los dispositivos concretos, la literatura sobre monitorización remota de salud coincide en señalar tres factores críticos para el éxito de este tipo de soluciones. El primero es la adherencia: de nada sirve un sistema preciso si el usuario deja de llevar el dispositivo. El segundo es la gestión de las falsas alarmas, que erosionan la confianza y conducen al abandono. El tercero es la integración del dato en una cadena de actuación clara, de modo que la detección desemboque en una respuesta útil. Kora aborda estos tres factores desde el diseño: un dispositivo cómodo para favorecer la adherencia, una clasificación de alertas que minimiza las falsas alarmas, y un flujo de notificación que conecta la detección con la respuesta de la familia.

2.3. Detección de anomalías en señales biométricas

La detección de anomalías es un campo amplio que abarca desde técnicas estadísticas clásicas hasta modelos de aprendizaje profundo. Para el alcance de este trabajo conviene distinguir tres familias.

Métodos estadísticos univariantes. Comparan cada medida con una distribución de referencia y marcan como anómalos los valores que caen fuera de un intervalo predefinido. Son simples, interpretables y baratos, pero no capturan relaciones entre variables ni tendencias temporales. La regla "frecuencia cardíaca en reposo por encima de 100 latidos por minuto" pertenece a esta familia.

Métodos basados en aislamiento o distancia. Identifican los puntos que se separan del resto de la nube de datos. Isolation Forest, la técnica elegida en este trabajo, pertenece a esta familia. Funciona bien cuando se dispone de suficientes datos históricos del propio usuario para construir un patrón de normalidad personalizado. Se ha elegido por tres motivos: es interpretable, es ligera en cómputo y está bien soportada por la librería scikit-learn, ampliamente utilizada en el grado.

Métodos basados en series temporales. Modelos como los autoencoders recurrentes o los modelos probabilísticos capturan dependencias temporales. Son los más potentes, pero también los más exigentes en datos y cómputo. Constituyen el horizonte natural de Kora cuando se disponga de una base de usuarios suficiente, pero quedan fuera del alcance de este trabajo. Empezar por Isolation Forest es una decisión deliberada: la complejidad adicional no se justifica hasta que los datos la exigen.

2.4. Exposición de modelos mediante interfaces de programación

Un modelo de análisis solo aporta valor si puede integrarse en un producto. La forma habitual de lograrlo es exponerlo a través de una interfaz de programación de aplicaciones (API), que permite que otros componentes, como una aplicación móvil o un dispositivo, le envíen datos y reciban resultados. En el ecosistema Python, una de las opciones más sencillas y documentadas para este fin es el microframework Flask, que permite definir un servicio web con pocas líneas de código y resulta accesible para perfiles que no son ingenieros de software.

La literatura técnica reciente reconoce además el papel de las herramientas no-code, como n8n, Make o Zapier, para resolver tareas de automatización y orquestación sin necesidad de programar. Estas plataformas permiten definir flujos de trabajo mediante una interfaz visual y son especialmente adecuadas para problemas como el envío de notificaciones por distintos canales, donde el reto está en la configuración del flujo y no en la innovación algorítmica.

3. Análisis del entorno y oportunidad

3.1. Contextualización del problema

El análisis parte de tres datos difíciles de discutir. España cuenta con 9,7 millones de personas mayores de 65 años. Más de 1,7 millones de mayores de 70 viven solas. Y el sistema de cuidado, formado por residencias, centros de día y ayuda a domicilio, está saturado, con listas de espera de meses en muchas comunidades y un coste medio de plaza residencial en torno a los 2.000 euros mensuales, fuera del alcance de buena parte de las familias.

Para contrastar estas cifras con la percepción real, se realizó un cuestionario a 103 personas durante el primer trimestre del curso. Los resultados confirmaron tres aspectos. El primero, el carácter transversal del problema: el 73% de los encuestados con padres mayores de 70 años manifestó preocupación frecuente por su bienestar. El segundo, la escasez de información: el 81% reconoció no disponer de medios para conocer en tiempo real el estado de su familiar, más allá de llamarle. El tercero, la disposición a pagar: el 64% se mostró dispuesto a abonar entre 25 y 50 euros mensuales por una solución que aportara tranquilidad, siempre que fuera sencilla para el portador.

Una muestra de 103 respuestas no permite extrapolar con rigor estadístico al conjunto de la población, pero sí ofrece una señal coherente con lo que la literatura y los datos de mercado apuntan, y resulta suficiente para orientar las decisiones de una fase temprana de validación.

3.2. Análisis PESTEL

El análisis PESTEL se ha aplicado atendiendo a los factores que afectan de verdad a una solución como Kora, evitando el formalismo de rellenar todas las casillas sin criterio.

Factores políticos y legales

El cuidado de mayores ocupa un lugar creciente en la agenda política española y europea. El Plan de Recuperación incluye fondos específicos para la modernización del sistema de cuidados, y la Estrategia Nacional de Salud Digital apuesta por la teleasistencia avanzada. En el plano legal, el factor más relevante es el tratamiento de datos de salud, que el Reglamento General de Protección de Datos clasifica como categoría especialmente protegida, con exigencias técnicas elevadas de cifrado, control de accesos y registro de auditoría. Existe además un debate sobre si un sistema como Kora debe considerarse producto sanitario; la interpretación adoptada, respaldada por el asesoramiento legal del proyecto, es que mientras el

sistema se limite a detectar desviaciones y avisar, sin diagnosticar ni recomendar tratamiento, no encaja en la definición estricta de producto sanitario.

Factores económicos y sociales

El gasto sanitario público crece de forma estructural con el envejecimiento, lo que empuja hacia soluciones que reduzcan hospitalizaciones evitables. Por el lado de la demanda, el segmento de hijos de 45 a 60 años dispone de capacidad económica suficiente para una cuota mensual moderada, sobre todo si la propuesta se posiciona como prevención y tranquilidad. Dos tendencias sociales refuerzan la oportunidad: la dispersión geográfica de las familias y la mayor familiaridad digital de las generaciones que se acercan ahora a los 65 años.

Factores tecnológicos

El estado actual de los sensores, las baterías y las técnicas de análisis de datos permite construir la solución con componentes existentes y bien documentados. El reto no consiste en inventar tecnología nueva, sino en integrar con criterio piezas que ya están disponibles. Esta es precisamente la oportunidad para un perfil de análisis de negocios: aportar el componente analítico que convierte datos en decisiones útiles.

Conviene matizar, no obstante, que la disponibilidad de la tecnología no garantiza por sí sola el éxito. El reto real no es técnico, sino de integración y de modelo de negocio: combinar componentes existentes de forma que el resultado sea fiable, asequible y fácil de usar para un público que no es nativo digital. Esta observación refuerza la pertinencia de abordar el proyecto desde un perfil de análisis de negocios, capaz de tender el puente entre la capacidad técnica y la necesidad del mercado.

3.3. Análisis competitivo y posicionamiento

El mapa competitivo se organiza en tres anillos. El más cercano lo forman las soluciones de teleasistencia y eldercare digital. En España destacan Durcal, que comercializa un dispositivo similar a un reloj con botón de emergencia y detección de caídas por una cuota mensual de 25 a 35 euros; Atenzia, que opera el servicio público de teleasistencia en varias comunidades; y Saralud, posicionada en el cruce sanitario. Ninguna ha apostado de forma decidida por la prevención basada en el análisis de patrones: el modelo sigue siendo reactivo.

El anillo intermedio lo forman los wearables generalistas, técnicamente solventes pero no diseñados para el segmento, con interfaces que exigen cierta competencia digital y algoritmos no entrenados con población mayor. El anillo más amplio lo forman los servicios tradicionales,

como residencias y cuidadores a domicilio, que no resuelven el mismo problema pero compiten por el presupuesto de la familia.

El posicionamiento de Kora no es de confrontación con ninguno de los tres anillos, sino de ocupación de un espacio que hoy nadie cubre con claridad: la prevención basada en el análisis del patrón individual, asequible para la familia y diseñada para integrarse en el día a día sin exigir nada al usuario. Ese espacio es defendible porque requiere a la vez tres cosas que los competidores actuales no reúnen: hardware adaptado, un modelo analítico entrenado con la población diana y una experiencia de uso cuidada.

4. Modelo de negocio

Este capítulo encuadra el modelo de negocio de Kora con el fin de ofrecer el contexto necesario para comprender las decisiones analíticas y técnicas de los capítulos siguientes. El modelo de detección no se diseña en el vacío, sino al servicio de un producto con un público y una lógica de ingresos concretos.

4.1. Propuesta de valor

La propuesta de valor de Kora se resume en tres ideas. La primera es que está pensado desde el principio para personas mayores, con un dispositivo ergonómico y un modelo analítico entrenado con el patrón de la población diana, no con el de usuarios jóvenes y activos. La segunda es que anticipa en lugar de reaccionar: analiza tendencias a lo largo del tiempo para detectar señales de deterioro antes de que se conviertan en urgencias. La tercera es que ofrece tranquilidad a la familia y autonomía al mayor a la vez, dos beneficios que suelen presentarse como contradictorios y que aquí se concilian porque la familia recibe información y alertas mientras el mayor sigue viviendo como siempre.

Esta conciliación entre tranquilidad y autonomía es, en realidad, el núcleo de la propuesta. Las soluciones de vigilancia tradicionales tienden a resolver la tranquilidad de la familia a costa de la sensación de control del mayor, que percibe el dispositivo como una intromisión. Kora invierte esa lógica: el mayor apenas interactúa con la tecnología, que se limita a un dispositivo cómodo, mientras que la inteligencia del sistema trabaja en segundo plano y solo se manifiesta cuando hay algo relevante que comunicar. De este modo, el producto no pide al usuario que cambie sus hábitos ni que aprenda a manejar una interfaz, lo que elimina la principal barrera de adopción de la tecnología en este segmento.

4.2. Segmentos de cliente

Kora distingue entre usuario y comprador. El usuario es la persona mayor de 65 años, independiente pero con algún riesgo moderado, que vive sola o con su pareja. El comprador es el familiar, por lo general un hijo de entre 35 y 55 años, con poder adquisitivo medio-alto, que reside lejos de su familiar mayor y experimenta preocupación recurrente por su bienestar. Esta distinción es relevante para el diseño porque el sistema debe servir sobre todo al familiar, que es quien consulta la información y toma decisiones, manteniendo al mismo tiempo la mínima fricción posible para el mayor.

4.3. Fuentes de ingresos y estructura de costes

El modelo de negocio descansa sobre dos fuentes de ingresos. La primera, y principal en la fase de validación, es la venta a residencias geriátricas como herramienta de apoyo para los cuidadores, con una cuota que combina dispositivo en alquiler y suscripción mensual por residente. La segunda, que se activa tras la validación, es la suscripción de familias, que incluye dispositivo, monitorización y análisis, con un precio objetivo de entre 29 y 39 euros mensuales.

Los costes se reparten entre fijos, como sueldos, mantenimiento e infraestructura, y variables, como la fabricación del dispositivo, la logística y el marketing. Para el presente trabajo, lo relevante de esa estructura es que el coste de la parte analítica por usuario debe ser muy bajo para que el modelo funcione. Esta restricción condiciona las decisiones técnicas de los capítulos siguientes, orientadas hacia modelos ligeros que no requieren hardware especializado.

4.4. Encaje con el perfil del grado

La propuesta encaja con el perfil del Grado en Análisis de Negocios por una razón concreta. La disciplina atraviesa tres planos que tradicionalmente se abordan por separado: identificar y dimensionar un problema real, tratar y modelar los datos que lo describen, y traducir las conclusiones en decisiones empresariales. Kora ofrece la oportunidad de ejercitar los tres sobre un mismo caso. El componente que este trabajo desarrolla en profundidad, la detección de anomalías sobre datos biométricos, es análisis de datos aplicado en su sentido más estricto, y constituye la aportación técnica que la Guía Docente exige a un trabajo de esta naturaleza.

5. Datos y metodología analítica

5.1. Origen y naturaleza de los datos

El modelo de detección trabaja sobre cinco señales biométricas que un dispositivo vestible puede medir de forma continua: frecuencia cardíaca, saturación de oxígeno en sangre, temperatura corporal, número de pasos y horas de sueño. Cada medida lleva asociada una marca temporal que permite reconstruir la evolución de la señal a lo largo del día y de las semanas.

En la fase actual del proyecto no se dispone de datos reales de un dispositivo, por lo que el modelo se ha entrenado y validado sobre un conjunto de datos simulado. Esta es una limitación que se reconoce abiertamente y que se aborda en el capítulo de evaluación. El conjunto simulado se ha diseñado para reproducir patrones fisiológicos realistas y para incorporar anomalías controladas, lo que permite comprobar de forma objetiva si el modelo las detecta.

El uso de datos simulados en una fase temprana es una práctica habitual y metodológicamente válida cuando no se dispone todavía de datos reales. Presenta dos ventajas claras. La primera es que se conoce con certeza dónde están las anomalías, lo que permite medir el rendimiento del modelo de forma objetiva, algo imposible con datos reales sin etiquetar. La segunda es que evita por completo los problemas de privacidad y cumplimiento normativo asociados al tratamiento de datos de salud reales, que en esta fase del proyecto añadirían una complejidad innecesaria. La contrapartida, que se asume y se discute en el capítulo de limitaciones, es que los patrones simulados son más limpios que los reales.

5.2. Generación del conjunto simulado

El conjunto de datos se ha generado mediante un script en Python que reproduce el comportamiento de cada señal con criterios extraídos de la literatura fisiológica. La frecuencia cardíaca sigue un patrón circadiano, con valores más bajos durante la noche y más altos a media tarde, alrededor de una media individual. La saturación de oxígeno se mantiene estable con ruido moderado. La temperatura presenta una ligera oscilación diaria. Los pasos se concentran en las horas de actividad y son nulos durante la noche. Las horas de sueño se registran una vez al día.

La frecuencia de muestreo es otra decisión relevante del diseño. Las señales continuas, como la frecuencia cardíaca o la saturación, se registran cada pocos minutos, mientras que otras, como las horas de sueño, se resumen una vez al día. Este muestreo heterogéneo refleja la naturaleza de cada variable y condiciona la forma en que el modelo las trata. Un muestreo

demasiado frecuente generaría un volumen de datos innecesario y aumentaría el coste de almacenamiento sin aportar información relevante; uno demasiado escaso podría pasar por alto episodios breves pero significativos. El intervalo elegido busca el equilibrio entre capacidad de detección y eficiencia.

Sobre este patrón de normalidad se han inyectado anomalías controladas en días concretos: episodios de taquicardia sostenida, descensos nocturnos de la saturación, fiebre leve, caídas marcadas de la actividad y fragmentación del sueño. Conocer de antemano dónde están las anomalías permite medir con precisión la capacidad de detección del modelo, calculando cuántas identifica correctamente y cuántas marca por error.

5.3. Análisis exploratorio

Antes de modelar, se ha realizado un análisis exploratorio de las señales generadas con las herramientas habituales de la librería pandas y visualizaciones con matplotlib. El análisis confirma que las distribuciones de cada variable son coherentes con los rangos fisiológicos esperados y que las anomalías inyectadas se separan visiblemente del patrón habitual cuando se representan a lo largo del tiempo. Este paso, además de validar la calidad del conjunto simulado, orienta la elección de las variables que alimentan el modelo.

El análisis exploratorio cumple además una función de control de calidad de los propios datos. Antes de alimentar cualquier modelo, conviene comprobar que las señales no contienen valores imposibles, lagunas injustificadas o saltos que delaten un error en la generación. En el caso del conjunto simulado, la inspección de las distribuciones permitió confirmar que la frecuencia cardíaca se mantiene en rangos plausibles, que la saturación de oxígeno rara vez baja de valores normales salvo en las anomalías inyectadas, y que la actividad sigue el ciclo día-noche esperado. Esta verificación previa es una práctica habitual del análisis de datos y evita que errores de los datos se confundan después con hallazgos del modelo.

5.4. Construcción de la línea base individual

Una decisión central del trabajo es que la normalidad se define para cada usuario por separado, no por comparación con una media poblacional. Una frecuencia cardíaca en reposo de 75 latidos por minuto puede ser normal en una persona y anómala en otra cuyo reposo se sitúa de forma estable en torno a los 62. Por ello, para cada usuario se calcula una línea base a partir de sus medidas de las últimas semanas, que recoge la media, la desviación típica y los percentiles de cada señal. Esta línea base es la referencia individual contra la que se evalúan las medidas

nuevas, y se recalcula periódicamente para adaptarse a cambios sostenidos en el estado del usuario.

Antes de calcular nada conviene depurar y preparar los datos. La línea base se levanta descartando previamente las medidas ya marcadas como anómalas, para que la referencia no quede contaminada por los propios episodios que se quieren detectar. Después, todas las señales se estandarizan, restando la media y dividiendo por la desviación típica de cada variable. Este paso no es accesorio: la frecuencia cardíaca se mueve en torno a las decenas, los pasos diarios en miles y las horas de sueño en una sola cifra. Sin estandarizar, las variables de magnitud grande dominarían el cálculo de distancias simplemente por el tamaño de sus números, y no porque aporten más información. Al llevar todas las señales a una escala común, cada una contribuye en igualdad de condiciones a definir qué es lo habitual para esa persona.

El recálculo periódico de la línea base merece una reflexión. Un cambio sostenido en el patrón de un usuario puede deberse a dos causas opuestas: una evolución natural y benigna, como una mejora en la condición física, o un deterioro progresivo que el sistema debería señalar. Distinguir entre ambas no es trivial. La solución adoptada consiste en recalcular la línea base con una ventana móvil suficientemente larga para absorber la variabilidad normal, pero acompañada de un seguimiento de la tendencia a medio plazo. De este modo, un deterioro lento no queda enmascarado por la propia actualización de la referencia. Este es uno de los puntos donde el criterio analítico resulta más determinante, y donde la validación con datos reales aportará el mayor aprendizaje.

6. Modelo de detección de anomalías

6.1. Elección del algoritmo: Isolation Forest

El modelo de detección es el centro de este trabajo. Tras valorar las tres familias de técnicas descritas en el marco teórico, se ha elegido Isolation Forest, un algoritmo de la librería scikit-learn que detecta anomalías aislando observaciones. La idea del algoritmo es intuitiva: una observación anómala se separa del resto con pocas particiones aleatorias, mientras que una observación normal, situada en una zona densa de datos, requiere muchas más. Esa facilidad de aislamiento es la que el algoritmo traduce en una puntuación de anomalía.

En términos algo más concretos, Isolation Forest construye un conjunto de árboles de decisión en los que, en cada nodo, se elige al azar una variable y un valor de corte. Cada observación recorre los árboles hasta quedar aislada en una hoja. Las observaciones anómalas, al estar en zonas poco pobladas del espacio de datos, tienden a aislarse con un número reducido de cortes, mientras que las normales necesitan más. La puntuación de anomalía de cada observación se calcula a partir de la profundidad media a la que queda aislada en el conjunto de árboles. El parámetro de contaminación, que indica la proporción esperada de anomalías, permite ajustar el umbral a partir del cual una observación se considera anómala; en este trabajo se ha fijado en un valor bajo, coherente con el hecho de que las anomalías fisiológicas son, por definición, poco frecuentes.

La elección responde a tres criterios. El primero es la interpretabilidad: se puede explicar por qué una observación se considera anómala, lo cual es importante para generar confianza en las familias. El segundo es la ligereza: el algoritmo entrena y predice en milisegundos sobre un procesador convencional, sin necesidad de hardware especializado, lo que mantiene bajo el coste por usuario. El tercero es la madurez de la herramienta, ampliamente documentada y utilizada en el grado.

6.2. Variables y entrenamiento

Para cada usuario con suficiente historial se entrena un modelo Isolation Forest propio. Las variables que lo alimentan son el conjunto de señales biométricas que el dispositivo mide de forma continua: frecuencia cardíaca, pasos diarios, horas y calidad del sueño, nivel de estrés, actividad física y presión arterial. Una decisión importante es que el modelo no examina cada señal por separado, sino todas a la vez. Esto le permite reconocer combinaciones anómalas que ninguna variable revelaría en solitario: una frecuencia cardíaca algo elevada puede ser normal, pero acompañada de pocas horas de sueño, presión alta y baja actividad deja de serlo. Es

justamente ese tipo de coincidencia entre señales lo que el enfoque multivariante captura y lo que una regla aislada sobre una sola medida pasaría por alto.

El entrenamiento utiliza únicamente registros normales, de modo que el modelo aprende cómo es el comportamiento habitual del usuario y reserva su capacidad de extrañeza para todo lo que se aleje de él. Sobre esos datos se construye un bosque de árboles de aislamiento; en la configuración empleada, ciento cincuenta árboles, un número que ofrece estimaciones estables sin un coste de cómputo apreciable. El parámetro de contaminación, que indica la proporción esperada de anomalías, se ha fijado en torno al cinco por ciento, coherente con el hecho de que las desviaciones fisiológicas son, por definición, poco frecuentes. La semilla de aleatoriedad se mantiene fija para que los resultados sean reproducibles, algo necesario tanto para validar el trabajo como para poder reentrenar el modelo y comparar versiones.

Esta arquitectura individualizada plantea un problema evidente: qué hacer con un usuario que acaba de empezar y todavía no tiene historial suficiente. Para ese caso, el sistema mantiene en paralelo un modelo global, entrenado con los datos normales de todos los usuarios. Cuando una persona no alcanza el mínimo de muestras necesario para tener un modelo propio, sus medidas se evalúan contra ese modelo global, que representa un patrón fisiológico promedio razonable. A medida que acumula registros, el sistema pasa a usar su modelo individual, mucho más afinado. La transición es silenciosa: el usuario nunca percibe el cambio, pero la precisión de la detección mejora conforme el sistema lo conoce mejor.

Cuando llega una medida nueva, el modelo correspondiente le asigna una puntuación de anomalía. Si esa puntuación cae por debajo de un umbral, la medida se marca como anómala y pasa a la fase de clasificación de gravedad. A este mecanismo de aprendizaje se le suma una última red de seguridad que no depende de ningún modelo: un conjunto de reglas absolutas basadas en rangos fisiológicos de seguridad. Una saturación de oxígeno peligrosamente baja o una frecuencia cardíaca en valores extremos se marcan como alerta con independencia del historial del usuario y de lo que diga el modelo.

La combinación de un modelo de aprendizaje no supervisado con un conjunto de reglas absolutas responde a una lógica de prudencia. El modelo aporta la capacidad de detectar desviaciones sutiles y personalizadas, pero necesita tiempo para aprender el patrón de cada usuario. Las reglas absolutas, en cambio, no requieren aprendizaje y cubren las situaciones de riesgo evidente desde el primer momento. Este doble mecanismo garantiza que el sistema sea

seguro desde el primer día y que vaya afinando su sensibilidad a medida que acumula datos del usuario.

6.3. Clasificación por niveles de gravedad

No todas las anomalías tienen la misma importancia, y notificar todas por igual destruiría la confianza en el sistema, porque un sistema que avisa de todo acaba siendo un sistema que se ignora. Por ello, las anomalías detectadas se clasifican en tres niveles. El nivel informativo recoge desviaciones leves que el sistema registra pero no comunica, y que sirven para afinar el patrón. El nivel de atención corresponde a desviaciones sostenidas que la familia debería conocer pero que no exigen actuar de inmediato, y que se agrupan en un resumen diario. El nivel de urgencia se reserva para desviaciones agudas o combinaciones que podrían indicar un problema inminente, y que generan un aviso inmediato.

El criterio para situar una medida en uno u otro nivel no es arbitrario. El Isolation Forest no devuelve solo un veredicto de normal o anómalo, sino una puntuación continua que mide cuánto se aleja una observación del comportamiento aprendido: cuanto más extrema es, más baja es su puntuación. Los umbrales que separan los tres niveles se han calibrado de forma empírica observando la distribución de esas puntuaciones sobre los datos normales. El nivel más leve se corresponde aproximadamente con el cinco por ciento de medidas más extremas dentro de lo que aún se considera normal; el nivel intermedio, con alrededor del dos por ciento; y el nivel más grave, con apenas el uno por ciento, valores que prácticamente nunca aparecen en un comportamiento sano. Es decir, los umbrales no se han fijado de oído, sino anclados a los percentiles reales de la distribución de puntuaciones.

A esta clasificación por puntuación se superpone la capa de reglas absolutas. Una señal que cruza un umbral fisiológico crítico se eleva directamente a urgencia, por moderada que sea su puntuación de anomalía, de modo que un valor claramente peligroso nunca queda relegado a un simple resumen diario. El diseño concreto de dónde colocar las fronteras entre niveles es, probablemente, el aspecto que más ajuste requerirá cuando se disponga de datos reales y de la respuesta de usuarios reales: situarlas demasiado abajo genera fatiga por exceso de avisos, situarlas demasiado arriba deja pasar problemas. Así se reconoce en el capítulo de limitaciones.

6.4. Evaluación del modelo

Para validar el modelo se ha generado un conjunto simulado de un usuario a lo largo de sesenta días, con cincuenta anomalías inyectadas de tipos diversos. El modelo detectó cuarenta y cuatro de ellas, lo que supone una sensibilidad del 88%, y generó siete falsos positivos sobre las

medidas normales, lo que supone una especificidad del 99,4%. Estos resultados son razonables para un prototipo, si bien deben interpretarse con cautela: las anomalías reales son más sutiles que las simuladas, por lo que cabe esperar que la sensibilidad y la especificidad reales sean algo inferiores. La evaluación se detalla en el capítulo 9.

Para interpretar correctamente estas cifras conviene recordar el compromiso entre sensibilidad y especificidad. Una sensibilidad alta significa que el modelo detecta la mayoría de las anomalías reales, lo que es deseable en un contexto de cuidado, donde dejar pasar un problema tiene un coste elevado. Una especificidad alta significa que el modelo no genera demasiadas falsas alarmas, lo que es igualmente importante porque un exceso de avisos erosiona la confianza de la familia. El ajuste del umbral de detección desplaza el equilibrio entre ambos objetivos, y la elección del punto adecuado depende de cuánto pese cada tipo de error en el contexto concreto de uso.

En el caso de Kora, el coste de los dos tipos de error no es simétrico. Un falso negativo, es decir, no detectar una anomalía real, puede tener consecuencias graves para la salud del usuario. Un falso positivo, en cambio, supone una molestia y una pérdida de confianza, pero no un daño directo. Esta asimetría sugiere inclinar el ajuste hacia una mayor sensibilidad, aceptando algunos falsos positivos a cambio de no dejar pasar episodios reales. No obstante, esta inclinación tiene un límite: si los falsos positivos se vuelven demasiado frecuentes, la familia deja de atender las alertas y el sistema pierde su utilidad. El punto de equilibrio óptimo es, por tanto, una decisión que combina criterio analítico y conocimiento del contexto de uso, y que deberá refinarse con la experiencia real de los usuarios.

7. Exposición del modelo: la API de monitorización

7.1. Propósito y alcance

El modelo de detección solo resulta útil si otros componentes del sistema pueden enviarle datos y recibir resultados. Para ello se ha construido una interfaz de programación sencilla con el microframework Flask de Python. La elección de Flask, frente a alternativas más complejas, responde al criterio de mantener el desarrollo dentro de lo que es comprensible y defendible para un perfil de análisis de datos: Flask permite definir un servicio web con muy pocas líneas de código y es una de las herramientas más documentadas del ecosistema Python.

El alcance de la interfaz es deliberadamente reducido. Recibe medidas biométricas, las almacena, ejecuta sobre ellas el modelo de detección y devuelve el estado del usuario y las alertas generadas. No incluye funcionalidades de autenticación avanzada ni una arquitectura compleja, que quedan fuera del alcance de un trabajo de análisis de negocios y se mencionan únicamente como líneas futuras.

7.2. Endpoints y flujo de datos

La interfaz expone un conjunto reducido de puntos de acceso. Un punto de acceso recibe un lote de medidas biométricas, las almacena y dispara el análisis de anomalías. Otro devuelve el estado actual del usuario, entendido como el nivel de alerta más alto registrado en las últimas horas. Un tercero devuelve el histórico de alertas, y un cuarto permite recalcular la línea base del usuario. El flujo típico es sencillo: el dispositivo envía medidas cada cierto intervalo, la interfaz las procesa, y si detecta una anomalía de urgencia, comunica el evento al sistema de notificaciones descrito en el capítulo siguiente.

Los datos se intercambian en formato JSON, un estándar legible y ampliamente utilizado, y se almacenan en una base de datos sencilla. Cada punto de acceso tiene un contrato claro sobre qué información espera recibir y qué devuelve, lo que reduce los errores cuando otros componentes se comunican con la interfaz.

7.3. Decisiones de diseño

Dos decisiones de diseño merecen mención. La primera es la separación entre el modelo de análisis y la interfaz que lo expone: el modelo vive en su propio módulo y la interfaz se limita a llamarlo. Esta separación permite mejorar o sustituir el modelo sin tocar la interfaz, algo valioso en un proyecto donde la parte analítica evolucionará más que el resto. La segunda es la

simplicidad: en cada decisión se ha preferido la opción más sencilla y documentada que cumpliera el objetivo, evitando introducir complejidad que no aportara valor en esta fase.

Esta preferencia por la simplicidad no es una limitación, sino un principio de diseño consciente. En un proyecto en fase temprana de validación, la prioridad es comprobar si la propuesta de valor resuelve un problema real, no construir una infraestructura preparada para millones de usuarios que quizá nunca lleguen. Invertir esfuerzo en una arquitectura sofisticada antes de validar el producto sería una mala asignación de recursos. El sistema está diseñado para ser suficiente en la fase actual y para poder evolucionar cuando la tracción lo justifique, momento en el que las funcionalidades de seguridad, escalabilidad y rendimiento que ahora quedan fuera de alcance pasarían a primer plano.

8. Gestión de alertas con herramientas no-code

8.1. Justificación del enfoque no-code

El envío de alertas a los familiares es una tarea distinta del análisis de datos. No requiere desarrollar un algoritmo, sino orquestar un flujo: recibir un evento, decidir a quién avisar y por qué canal, y enviar el mensaje. Este tipo de problema se resuelve de forma más eficiente con una herramienta de automatización no-code que programándolo desde cero. Para este trabajo se ha empleado n8n, una plataforma de código abierto que permite definir flujos mediante una interfaz visual y que puede alojarse en el propio servidor, lo que ayuda a mantener el control sobre los datos.

La elección del enfoque no-code es una decisión consciente y defendible. Permite construir un sistema de notificaciones funcional sin necesidad de conocimientos avanzados de programación, concentra el esfuerzo técnico del trabajo en la parte que de verdad aporta valor analítico, y resulta fácil de mantener y modificar, ya que añadir un canal nuevo de aviso consiste en arrastrar un elemento en la interfaz.

8.2. Diseño del flujo de notificaciones

El flujo diseñado en n8n tiene cuatro fases. En la primera, recibe el evento que la interfaz de monitorización le envía cuando detecta una anomalía relevante. En la segunda, recupera la información de contacto y las preferencias del familiar. En la tercera, decide qué hacer según el nivel de la alerta. En la cuarta, envía el mensaje por el canal correspondiente y registra el resultado. Todo el flujo se define visualmente, sin escribir código, lo que facilita su comprensión y su mantenimiento.

8.3. Reglas de escalado

Las reglas de escalado se han diseñado para evitar la fatiga por exceso de notificaciones, que es uno de los principales riesgos de los sistemas de alerta. Una alerta de urgencia se notifica primero al contacto principal por el canal más inmediato; si no hay confirmación en unos minutos, se intenta por un segundo canal y, en último término, se avisa a un contacto secundario. Una alerta de atención no genera aviso inmediato, sino que se acumula en un resumen diario. Una alerta informativa solo se registra. Estas reglas son configurables para adaptarse a las preferencias de cada familia.

La lógica de escalado refleja un principio que atraviesa todo el diseño de Kora: la notificación es un recurso escaso que debe administrarse con cuidado. Cada aviso que llega a la familia

consume una parte de su atención y de su confianza en el sistema. Un aviso oportuno y bien fundamentado refuerza esa confianza; una sucesión de avisos triviales la destruye. Por eso la clasificación de las alertas y las reglas de escalado no son un detalle accesorio, sino una parte central de la propuesta de valor, tan importante como la precisión del propio modelo de detección.

9. Evaluación de la solución

9.1. Resultados del modelo

La evaluación del modelo de detección se ha realizado sobre el conjunto simulado descrito en el capítulo 5. Sobre cincuenta anomalías inyectadas a lo largo de sesenta días, el modelo detectó cuarenta y cuatro, con una sensibilidad del 88%, y produjo siete falsos positivos, con una especificidad del 99,4%. La siguiente tabla resume los resultados por tipo de anomalía.

Tipo de anomalía	Inyectadas	Detectadas	Tasa
Taquicardia sostenida	12	12	100%
Desaturación nocturna	10	9	90%
Fiebre leve sostenida	10	9	90%
Caída de actividad	10	8	80%
Fragmentación del sueño	8	6	75%
Total	50	44	88%

Tabla 1. Resultados de detección por tipo de anomalía sobre el conjunto simulado. Elaboración propia.

Los tipos de anomalía más agudos y de mayor amplitud, como la taquicardia sostenida, se detectan con tasas cercanas al total, mientras que los más sutiles, como la fragmentación del sueño, presentan tasas algo inferiores. Es un comportamiento esperable y coherente con la naturaleza del algoritmo, que aísla con más facilidad las observaciones que más se separan del patrón.

Los siete falsos positivos registrados se concentraron en días con variabilidad legítima del comportamiento, como jornadas de mayor actividad física o noches de sueño más corto pero no patológico. Este resultado sugiere una vía de mejora para fases posteriores: incorporar al modelo información de contexto, como el día de la semana o eventos conocidos, que permita distinguir la variabilidad normal de la anómala. Es una línea coherente con la evolución natural del sistema hacia modelos que tengan más en cuenta la dimensión temporal.

9.2. Viabilidad económica

Un modelo que funciona no basta si su coste de operación hace inviable el negocio. La parte analítica de Kora es deliberadamente ligera: el modelo Isolation Forest se ejecuta sobre un procesador convencional, sin hardware especializado, y la interfaz que lo expone tiene unos

requisitos mínimos. La estimación de coste de infraestructura para distintos niveles de uso se resume en la siguiente tabla.

Usuarios activos	Coste mensual estimado	Coste por usuario
100	45 €	0,45 €
1.000	165 €	0,17 €
10.000	1.200 €	0,12 €

Tabla 2. Estimación de coste de infraestructura de la parte analítica por niveles de carga. Elaboración propia.

El coste de infraestructura por usuario es muy bajo, inferior a medio euro mensual incluso con cien usuarios, y desciende con la escala. Frente a un precio de suscripción objetivo de entre 29 y 39 euros, la parte analítica representa una fracción mínima del coste de servicio, lo que confirma que la solución es compatible con un modelo de negocio sostenible.

Este resultado tiene una implicación estratégica relevante. Dado que el coste marginal de analizar a un usuario adicional es muy reducido, el modelo de negocio escala favorablemente: a medida que crece la base de usuarios, el coste por usuario disminuye y el margen mejora. El grueso del coste de Kora no está en la parte analítica, sino en la fabricación del dispositivo, la logística y la captación de clientes. Esto orienta las prioridades del proyecto: la palanca de rentabilidad no es optimizar el modelo para ahorrar cómputo, sino reducir el coste del hardware y mejorar la eficiencia de la captación. El análisis de viabilidad económica, por tanto, no solo confirma que la solución es sostenible, sino que también señala dónde conviene concentrar los esfuerzos de mejora.

9.3. Limitaciones

Una evaluación honesta debe reconocer sus limitaciones. La principal es que el modelo se ha validado sobre datos simulados, no reales; aunque el conjunto reproduce patrones plausibles, el comportamiento con datos de personas reales será distinto, y solo podrá confirmarse en una fase de validación con dispositivos físicos. La segunda es que no existe todavía integración con un dispositivo real, por lo que la captura de datos se asume pero no se ha implementado. La tercera es que el diseño de los umbrales de gravedad requerirá ajuste con datos y usuarios reales. La cuarta es que la estimación económica se basa en precios públicos de proveedores y constituye una aproximación, no una cifra definitiva.

Junto a las limitaciones hay un riesgo regulatorio que conviene tener presente: si en algún momento el regulador considerase que Kora es producto sanitario en sentido estricto, el coste

de certificación sería significativo. Las medidas de protección de datos implementadas son sólidas, pero ningún sistema está completamente exento de riesgo.

10. Conclusiones

10.1. Cumplimiento de objetivos

El objetivo general del trabajo era diseñar, entrenar y evaluar un sistema de detección de anomalías sobre datos biométricos, y exponerlo de forma integrable en el ecosistema Kora. Ese objetivo se ha cumplido. El modelo está construido, se ha validado sobre un conjunto simulado con resultados razonables, se expone a través de una interfaz sencilla y el envío de alertas se resuelve con una herramienta no-code. Los objetivos específicos también se han alcanzado: el análisis del sector, el encuadre del modelo de negocio, la generación del conjunto de datos, la construcción y evaluación del modelo y la estimación económica están todos desarrollados.

10.2. Aportaciones

La aportación principal del trabajo se sitúa en el plano analítico. Demuestra que es posible construir un sistema de detección de anomalías para el cuidado de mayores con técnicas estándar, interpretables y de bajo coste, y que ese sistema alcanza una calidad suficiente para sostener una primera fase de validación. La solución combina un modelo individual por usuario con un modelo global de respaldo y una capa de reglas fisiológicas absolutas, un diseño que mantiene la detección segura desde el primer día y la afina a medida que conoce a cada persona. Lo que más he aprendido durante el proceso, y que merece nombrarse, es que el valor de un proyecto como Kora no está solo en el modelo ni solo en el plan de negocio, sino en la conexión entre ambos: un modelo sin encaje de mercado es un ejercicio académico, y un plan de negocio sin un producto real es una promesa vacía.

Empecé el grado entendiendo el análisis de negocios como un conjunto de herramientas para tratar datos. Terminé este trabajo con una idea más amplia: lo que distingue a la disciplina es la capacidad de cerrar el círculo completo, desde la identificación de un problema real hasta la decisión de negocio, pasando por el dato, el modelo y su puesta en funcionamiento. Kora ha sido, en mi caso, el primer proyecto en el que he visto cerrarse ese círculo de principio a fin.

10.3. Líneas futuras

El trabajo abre varias líneas de continuación. La primera es validar el modelo con datos reales recogidos en una fase piloto, lo que permitiría reevaluar la elección de Isolation Forest frente a modelos de series temporales más sofisticados. La segunda es integrar el sistema con un dispositivo físico real. La tercera es enriquecer la interfaz de programación con las funcionalidades de seguridad y escalabilidad propias de un entorno de producción, que en este trabajo se han dejado deliberadamente fuera de alcance. La cuarta es explorar la integración

con los sistemas de información sanitaria, lo que abriría una vía de colaboración con profesionales clínicos.

Más allá de estas líneas concretas, el trabajo deja abierta una reflexión de mayor alcance sobre el papel del análisis de datos en el cuidado de las personas. La tecnología no sustituye el cuidado humano, sino que puede liberar tiempo y atención para que ese cuidado se concentre donde más se necesita. Un sistema como Kora no pretende reemplazar la llamada de un hijo a su madre, sino ofrecer un fondo de tranquilidad que haga esa relación más libre y menos angustiada. Esta perspectiva, que sitúa la técnica al servicio de las personas, es la que ha guiado el conjunto del trabajo y la que, a juicio del autor, mejor define el sentido del análisis de negocios aplicado a un problema social.

DECLARACIÓN DE USO DE IA GENERATIVA

Enrique Quintero Sieira, autor del Trabajo Fin de Grado titulado “PROYECTO KORA: DISEÑO Y DESARROLLO DE APIs PARA LA MONITORIZACIÓN DE MAYORES”, declaro de forma responsable que:

He utilizado herramientas de inteligencia artificial generativa durante la elaboración de este trabajo de manera responsable, transparente y como apoyo al proceso de investigación y redacción, sin que ello sustituya en ningún caso mi propio criterio, análisis y autoría.

En concreto, declaro el siguiente uso:

- Herramientas empleadas: [p. ej., Claude (Anthropic), ChatGPT (OpenAI), etc.].
- Finalidad del uso: apoyo en la revisión y corrección lingüística y de estilo, reorganización y mejora de la redacción, asistencia en la depuración y documentación de código, consulta de conceptos técnicos y orientación estructural del documento.
- Alcance: las herramientas se han empleado como asistencia complementaria. La concepción del trabajo, el diseño metodológico, el análisis de los datos, los resultados, las conclusiones y las decisiones técnicas son de mi exclusiva autoría.

Asimismo, declaro que he verificado y contrastado toda la información generada o sugerida por estas herramientas, asumiendo la plena responsabilidad sobre el contenido final del trabajo, su originalidad y la veracidad de las fuentes citadas. Confirmando que el uso de estas herramientas se ajusta a la normativa de integridad académica de la Universidad Pontificia Comillas.

En [Madrid], a 16 de junio de 2026.

Firmado:

Enrique Quintero Sieira

Bibliografía

Comisión Europea (2024). The Silver Economy: A growth opportunity for Europe. DG GROW. Publications Office of the European Union.

Liu, F. T., Ting, K. M., y Zhou, Z. H. (2008). Isolation Forest. En 2008 Eighth IEEE International Conference on Data Mining (pp. 413-422). IEEE.

OECD (2024). Health at a Glance: Europe 2024. OECD Publishing.

Perez, M. V., Mahaffey, K. W., Hedlin, H., y otros (2019). Large-scale assessment of a smartwatch to identify atrial fibrillation. *New England Journal of Medicine*, 381(20), 1909-1917.

Pérez Díaz, J., Castillo Belmonte, A. B., Aceituno Nieto, P., y Ramiro Fariñas, D. (2024). Un perfil de las personas mayores en España, 2024: Indicadores estadísticos básicos. Consejo Superior de Investigaciones Científicas (CSIC).

Pedregosa, F., y otros (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.

Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales (RGPD).

Grinberg, M. (2018). *Flask Web Development: Developing Web Applications with Python* (2.^a ed.). O'Reilly Media.

McKinney, W. (2017). *Python for Data Analysis* (2.^a ed.). O'Reilly Media.

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44-56.

World Health Organization (2024). World report on ageing and health, 2024 update. WHO Press.

Anexos

Anexo I. Repositorio de código

Todo el código desarrollado para este trabajo está disponible en un repositorio de GitHub, que incluye el script de generación de datos simulados, el módulo del modelo de detección de anomalías, la interfaz de programación en Flask y la documentación del flujo de notificaciones de n8n. El repositorio permite reproducir los resultados de la evaluación.

URL del repositorio: [kikeiro71/tfg-anomalias](https://github.com/kikeiro71/tfg-anomalias)

Anexo II. Estructura del conjunto de datos

El conjunto de datos simulado se compone de registros con cuatro campos: identificador de usuario, tipo de señal, valor y marca temporal. El script de generación y el detalle de los parámetros fisiológicos empleados se encuentran documentados en el repositorio.

Anexo III. Flujo de notificaciones

El flujo de notificaciones diseñado en n8n se encuentra exportado en formato JSON en el repositorio, junto con las instrucciones para importarlo en una instancia de la herramienta.