



Diseño e implementación de una arquitectura de software basada en LLM para la detección y clasificación inteligente de emergencias multi escenario: Caso LivIApp

David Valdés Ponce

david.valdesp@usm.cl

Mauricio Figueroa
Profesor Guía

Resumen: LivIApp es una plataforma móvil orientada a la generación de alertas ante situaciones de riesgo. Su mecanismo inicial basado en palabras clave presenta limitaciones frente a la variabilidad del lenguaje humano y restringe su capacidad para abordar distintos tipos de emergencias. Este trabajo propone una arquitectura de software basada en modelos de lenguaje de gran escala para incorporar detección y clasificación semántica multi escenario.

La solución integra captura de audio, transcripción mediante Whisper y análisis con GPT-4o-mini, generando una respuesta estructurada que identifica la existencia, categoría y gravedad de una posible emergencia. Además, incorpora un mecanismo de doble verificación que utiliza contexto temporal posterior para confirmar la situación o descartar falsas alarmas antes de alertar a la red de apoyo.

La arquitectura se evaluó mediante 40 escenarios simulados de seguridad, emergencias médicas y accidentes. Bajo condiciones controladas, las decisiones finales coincidieron con el resultado esperado en los 40 casos. En aquellos que requirieron doble verificación, la latencia promedio acumulada de procesamiento posterior a la captura fue de 7,87 segundos, excluyendo la duración de las ventanas de grabación. Estos resultados proporcionan evidencia inicial sobre la viabilidad del análisis semántico basado en LLM.

Palabras Clave: LivIApp, modelos de lenguaje, detección de emergencias, procesamiento de lenguaje natural, reconocimiento automático del habla.

1 Introducción

1.1 Contexto

La incorporación de dispositivos móviles y tecnologías de inteligencia artificial ha ampliado las posibilidades de desarrollar sistemas capaces de apoyar la detección y comunicación de situaciones de emergencia. A diferencia de los mecanismos tradicionales, donde la solicitud de ayuda depende principalmente de una acción explícita del usuario, las soluciones actuales pueden utilizar distintas fuentes de información para identificar situaciones de riesgo y generar respuestas de manera automatizada.

Dentro de este contexto se encuentra LivIApp, una aplicación móvil concebida inicialmente como una herramienta de apoyo frente a situaciones de violencia intrafamiliar. La plataforma incorpora mecanismos de captura de audio, ubicación y comunicación con una red de personas de confianza, permitiendo generar alertas cuando se identifica una situación potencialmente peligrosa.

La versión inicial del sistema utiliza un mecanismo de reconocimiento basado en palabras clave. Si bien este enfoque permite una detección directa en escenarios previamente definidos, su capacidad de generalización es

CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☒ Memoria o trabajo de título ☐ Tesis de Postgrado

Título del trabajo: Diseño e implementación de una arquitectura de software basada en LLM para la detección y clasificación inteligente de emergencias multi escenario: Caso LivlApp

Nombre del candidato(a): David Alejandro Valdés Ponce

Carrera / Grado: Ingeniería Civil Informática

Campus: Casa Central **Departamento:** Informática

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Mauricio Figueroa Colarte, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por (**marcar una opción**):

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 19 de octubre de 2026 **Firma:**



Estudiante o Candidato(a):



Fecha: 19 de octubre de 2026 **Firma:**

DAVID ALEJANDRO VALDÉS PONCE
RUT: 20.848.760-4

limitada cuando las personas expresan una situación utilizando palabras diferentes, frases incompletas o referencias indirectas. Esta característica adquiere especial relevancia al considerar la evolución de LivIApp hacia una plataforma capaz de abordar múltiples tipos de emergencias.

1.2 Motivación

La motivación de este trabajo surge de la necesidad de ampliar la capacidad de interpretación de LivIApp sin depender exclusivamente de expresiones previamente configuradas. En una situación crítica, la forma en que una persona comunica lo que está ocurriendo puede variar significativamente dependiendo del contexto, su estado emocional y las características del incidente.

La utilización de técnicas de reconocimiento automático del habla y modelos de lenguaje permite abordar esta limitación mediante una separación entre la transcripción de la información acústica y su posterior interpretación semántica [1], [2]. Esto abre la posibilidad de identificar situaciones donde la emergencia no se manifiesta mediante una palabra específica, sino mediante la relación contextual entre diferentes elementos de una conversación.

La incorporación de estas capacidades también permite ampliar el alcance de LivIApp desde su escenario inicial hacia situaciones relacionadas con seguridad, emergencias médicas y accidentes, manteniendo la comunicación posterior con una red de apoyo.

1.3 Problema

El mecanismo de detección basado en palabras clave presenta una dependencia directa de un conjunto de expresiones previamente definidas [3]. Como consecuencia, una situación real puede no ser detectada cuando la persona utiliza una formulación diferente a la considerada durante la configuración del sistema.

Este problema se vuelve más significativo al incorporar múltiples categorías de emergencia. Situaciones de seguridad, accidentes domésticos y emergencias médicas poseen características lingüísticas diferentes y pueden ser descritas de múltiples maneras. Mantener manualmente un conjunto suficientemente representativo de términos para todos estos escenarios incrementa progresivamente la complejidad y dificulta la escalabilidad del mecanismo de detección.

Por otra parte, una interpretación automática excesivamente sensible puede generar falsas alarmas cuando determinadas expresiones aparecen en conversaciones cotidianas, referencias a eventos pasados o contenido audiovisual. Por lo tanto, el problema no consiste únicamente en detectar términos asociados a situaciones de riesgo, sino en determinar si el contexto de la comunicación representa efectivamente una emergencia.

1.4 Definición del Problema

A partir de las limitaciones identificadas, el problema abordado en esta memoria corresponde al diseño de una arquitectura capaz de transformar información de audio en una decisión estructurada sobre la existencia y características de una posible emergencia, considerando el contexto de la comunicación y permitiendo su utilización posterior por LivIApp.

La solución debe integrar mecanismos de reconocimiento de voz y análisis semántico, clasificar distintos tipos de emergencia, generar información comprensible para la red de apoyo y reducir la generación de falsas alarmas mediante mecanismos adicionales de verificación, manteniendo tiempos de procesamiento compatibles con el contexto de uso de la plataforma.



1.5 Propuesta de Solución y Objetivos

La solución propuesta consiste en reemplazar el enfoque basado exclusivamente en palabras clave por una arquitectura de procesamiento compuesta por distintas etapas. El audio capturado por LivIApp es transformado en texto mediante Whisper y posteriormente analizado por un modelo de lenguaje de gran escala, configurado para determinar si existe una posible emergencia y extraer información relacionada con su categoría y gravedad.

Cuando el primer análisis identifica una situación potencialmente riesgosa, la arquitectura incorpora una segunda etapa de verificación utilizando un segmento adicional de audio. La nueva información se analiza conjuntamente con el contexto inicial, permitiendo confirmar la emergencia o descartar una falsa alarma antes de generar la notificación definitiva.

Objetivo general

Diseñar, implementar y validar una arquitectura de software basada en modelos de lenguaje de gran escala para la detección y clasificación automática de situaciones de emergencia multi escenario a partir de información de audio, integrando transcripción, análisis semántico y mecanismos de verificación contextual en la plataforma LivIApp.

Objetivos específicos

- **Diseñar** e implementar un pipeline de procesamiento que integre la captura de audio, transcripción mediante Whisper y análisis semántico mediante un modelo de lenguaje, permitiendo transformar información de voz en una decisión estructurada utilizable por LivIApp.
- **Desarrollar** un motor de análisis semántico multi escenario capaz de identificar situaciones de seguridad, emergencias médicas y accidentes a partir de lenguaje natural no estructurado, determinando además la gravedad y el contexto asociado al incidente.
- **Implementar** un mecanismo de doble verificación y generación de alertas estructuradas, utilizando información contextual adicional para confirmar o descartar una sospecha inicial y generar un resumen de la situación destinado a la red de apoyo.
- **Evaluar** el funcionamiento de la solución en un ambiente controlado mediante escenarios simulados, considerando métricas de detección y clasificación, falsos positivos y falsos negativos, comportamiento del mecanismo de doble verificación y latencia de procesamiento posterior a la captura de audio, verificando que esta última se mantenga por debajo de 10 segundos bajo las condiciones de prueba establecidas.

1.6 Marco Teórico

El desarrollo de la solución requiere conceptos provenientes del reconocimiento automático del habla, procesamiento de lenguaje natural y modelos de lenguaje de gran escala. El reconocimiento automático del habla permite transformar la señal acústica en una representación textual, mientras que las técnicas de procesamiento de lenguaje natural permiten analizar posteriormente el significado y contexto de dicha información.

Sobre esta base, los modelos de lenguaje de gran escala proporcionan capacidades de análisis semántico y clasificación que permiten trabajar con lenguaje no estructurado sin depender exclusivamente de patrones predefinidos. Los fundamentos asociados a estas tecnologías y su relación con la arquitectura propuesta son desarrollados en el Capítulo 3.

1.7 Metodología de Validación

La solución se evalúa mediante una metodología experimental basada en escenarios simulados y ejecutados en un ambiente controlado. La validación se divide en dos baterías complementarias, alcanzando un total de 40

casos de prueba. La primera batería contempla 30 escenarios orientados a evaluar la detección y clasificación de situaciones de seguridad, emergencias médicas, accidentes y situaciones sin emergencia. La segunda incorpora 10 escenarios diseñados específicamente para analizar el comportamiento del mecanismo de doble verificación.

Los casos correspondientes a la validación final fueron ejecutados una vez finalizada la configuración del prompt utilizado por el modelo, sin ser empleados para modificar posteriormente sus instrucciones.

La evaluación considera métricas asociadas a verdaderos y falsos positivos y negativos, exactitud, precisión, recall y F1-score, además de la clasificación obtenida y los tiempos de procesamiento. La segunda batería permite también observar el efecto del contexto adicional sobre la confirmación de una emergencia y la estimación de su gravedad.

Debido al carácter controlado de los escenarios, los resultados obtenidos se utilizan como una evaluación funcional de la arquitectura implementada y no como una estimación de su desempeño frente a todas las condiciones posibles de una emergencia real.

1.8 Organización del Informe

El informe se organiza en seis capítulos. El Capítulo 1 introduce el contexto, motivación, problema, propuesta de solución, objetivos y metodología general del trabajo. El Capítulo 2 profundiza en la problemática abordada y presenta el estado del arte de soluciones relacionadas con la detección y comunicación de emergencias. El Capítulo 3 desarrolla los fundamentos conceptuales asociados al reconocimiento automático del habla, procesamiento de lenguaje natural y modelos de lenguaje de gran escala. El Capítulo 4 describe la arquitectura diseñada y su integración dentro de LivIApp, incluyendo el mecanismo de doble verificación. El Capítulo 5 presenta la metodología de validación y analiza los resultados obtenidos. Finalmente, el Capítulo 6 presenta las conclusiones, limitaciones y posibles líneas de trabajo futuro.

2 Definición del Problema y Estado del Arte

2.1 Análisis Profundo de la Problemática

El desarrollo de sistemas inteligentes para la detección de emergencias ha adquirido una importancia creciente en distintos ámbitos, como la seguridad ciudadana, la atención de emergencias médicas y la protección de personas en situación de vulnerabilidad. La incorporación de dispositivos móviles y tecnologías de inteligencia artificial ha permitido automatizar parte del proceso de detección y notificación de incidentes, con el objetivo de reducir los tiempos de respuesta y proporcionar información útil a quienes deben actuar frente a una situación de riesgo.

Sin embargo, la detección automática de una emergencia continúa siendo un desafío complejo. A diferencia de otros problemas de clasificación, estos escenarios involucran información incompleta, lenguaje espontáneo, condiciones de estrés y la necesidad de emitir una respuesta en tiempos reducidos. En consecuencia, la efectividad de estos sistemas no depende únicamente de reconocer una entrada de datos, sino también de interpretar correctamente el contexto en el que se produce y entregar información confiable para apoyar la toma de decisiones.

En el caso de LivIApp, esta problemática adquiere una relevancia particular debido a la necesidad de evolucionar desde un sistema orientado a un escenario específico y convertirlo en una plataforma capaz de responder a múltiples tipos de emergencias. Para comprender esta necesidad, resulta pertinente analizar las características de los enfoques tradicionales de detección, sus principales limitaciones y los desafíos que plantea



el procesamiento del lenguaje natural en contextos críticos, estableciendo finalmente los requerimientos que debe satisfacer una solución de nueva generación.

2.1.1 Sistemas tradicionales de detección de emergencias

Los sistemas de detección de emergencias han evolucionado significativamente en las últimas décadas como consecuencia del avance de las tecnologías de comunicación, el desarrollo de dispositivos móviles y la incorporación progresiva de herramientas de inteligencia artificial. Su principal objetivo consiste en identificar oportunamente situaciones que representen un riesgo para la integridad de las personas y facilitar una respuesta rápida por parte de los servicios de emergencia o de una red de apoyo.

En sus primeras implementaciones, estos sistemas se basaban principalmente en mecanismos de activación manual, tales como botones de pánico o llamadas telefónicas a servicios de emergencia. Si bien estas soluciones continúan siendo ampliamente utilizadas debido a su simplicidad y confiabilidad, presentan una limitación evidente: dependen completamente de que el usuario pueda solicitar ayuda de forma consciente y oportuna. En situaciones donde la persona pierde el conocimiento, se encuentra bajo coacción o posee movilidad reducida, este tipo de mecanismos puede resultar insuficiente.

Con el aumento de la capacidad de procesamiento de los dispositivos y la disponibilidad de sensores integrados, comenzaron a desarrollarse soluciones capaces de detectar automáticamente determinados eventos mediante acelerómetros, giroscopios, GPS, micrófonos y otros sensores [4]. Este enfoque permitió identificar situaciones como caídas, impactos o variaciones anómalas en parámetros fisiológicos, reduciendo la dependencia de la intervención directa del usuario. Sin embargo, la efectividad de estos sistemas suele estar restringida al tipo de sensor utilizado, dificultando la identificación de emergencias cuya principal fuente de información corresponde a la comunicación verbal o al contexto en que ocurre el incidente.

Más recientemente, la masificación de asistentes virtuales y tecnologías de reconocimiento de voz ha impulsado la incorporación del lenguaje hablado como una fuente adicional para la detección de emergencias. Este cambio ha permitido que los usuarios interactúen con los sistemas de manera más natural, eliminando parcialmente la necesidad de realizar acciones manuales para solicitar ayuda. No obstante, la incorporación de la voz también ha introducido nuevos desafíos relacionados con la interpretación del lenguaje humano, dando origen a enfoques que buscan identificar situaciones de riesgo a partir del contenido de las conversaciones.

En consecuencia, la evolución de los sistemas de detección de emergencias ha transitado desde mecanismos manuales hacia soluciones cada vez más automatizadas y orientadas al análisis de información contextual. Sin embargo, la forma en que estos sistemas interpretan el lenguaje continúa siendo un aspecto determinante para su efectividad, especialmente cuando deben operar en escenarios reales donde las personas no siguen patrones de comunicación predefinidos. Esta problemática será analizada en la siguiente subsección.

2.1.2 Limitaciones del reconocimiento basado en palabras clave

Uno de los enfoques más utilizados para la detección automática de emergencias mediante voz consiste en identificar palabras o expresiones previamente definidas que actúan como indicadores de una situación de riesgo [5]. Este tipo de soluciones se caracteriza por su simplicidad de implementación, baja carga computacional y tiempos de respuesta reducidos, razón por la cual ha sido ampliamente adoptado en aplicaciones donde se requiere una detección rápida y determinista.

En el caso de LivIApp, la versión inicial de la plataforma fue concebida para la detección de situaciones de violencia intrafamiliar mediante el reconocimiento de palabras clave asociadas a este tipo de incidentes. Este enfoque permitió desarrollar una solución funcional para un escenario específico, facilitando la activación de alertas cuando el usuario pronunciaba determinadas expresiones previamente configuradas. Sin embargo, al plantear la evolución de la plataforma hacia un sistema capaz de detectar múltiples tipos de emergencias,

comenzaron a evidenciarse limitaciones inherentes a este mecanismo de detección. Estas limitaciones fueron identificadas durante la definición del problema del proyecto, donde se establece que la dependencia de términos estáticos restringe la capacidad del sistema para interpretar situaciones reales y ampliar su cobertura hacia otros escenarios de riesgo.

La principal limitación de este enfoque radica en que el lenguaje humano es altamente variable. En una situación de emergencia, las personas no siguen necesariamente un conjunto predefinido de expresiones, sino que suelen comunicarse de forma espontánea, utilizando distintas palabras, frases incompletas o descripciones indirectas del incidente. Como consecuencia, un sistema basado exclusivamente en coincidencias de palabras puede no detectar correctamente una situación de riesgo cuando el usuario expresa la misma intención utilizando una formulación diferente [6].

2.1.3 Desafíos del procesamiento de lenguaje natural en escenarios críticos

La interpretación del lenguaje humano en escenarios críticos representa uno de los principales desafíos del procesamiento de lenguaje natural. A diferencia de otros dominios, las conversaciones asociadas a una emergencia suelen contener interrupciones, expresiones ambiguas, información incompleta y un alto componente emocional, factores que dificultan la correcta identificación de la intención del usuario [7]. Esta problemática adquiere mayor relevancia al considerar la ampliación del alcance funcional de LivIApp hacia un entorno multi escenario.

Escenarios como las emergencias médicas, accidentes domésticos o situaciones de inseguridad presentan características lingüísticas distintas entre sí y pueden describirse de múltiples maneras dependiendo del contexto, del estado emocional del usuario o incluso de la presencia de terceros durante la conversación.

En estas condiciones, mantener un conjunto de palabras clave lo suficientemente representativo resulta cada vez más complejo y difícil de escalar, incrementando tanto el riesgo de falsos negativos como el esfuerzo de mantenimiento del sistema.

2.1.4 Requerimientos de un sistema inteligente de respuesta

El análisis realizado permite establecer que un sistema moderno de detección inteligente de emergencias debe ser capaz de superar las limitaciones inherentes a los enfoques basados únicamente en reglas o palabras clave. Para ello, es necesario incorporar mecanismos que permitan comprender el contexto de una conversación, interpretar la intención del usuario y adaptarse a la diversidad de formas en que una emergencia puede ser expresada.

Además de la capacidad de interpretación semántica, este tipo de sistemas debe satisfacer una serie de requerimientos técnicos asociados a su utilización en escenarios críticos [8]. Entre ellos destacan la capacidad de operar con baja latencia, la escalabilidad para incorporar nuevos tipos de emergencias, la trazabilidad de las decisiones tomadas durante el procesamiento y la generación de información estructurada que facilite la toma de decisiones por parte de los equipos de respuesta.

Estos requerimientos constituyen la base para el diseño de una arquitectura de software capaz de integrar modelos avanzados de procesamiento del lenguaje natural dentro de un flujo confiable, escalable y orientado a aplicaciones donde el tiempo de respuesta y la calidad de la información son factores críticos.

2.2 Estado del Arte

El desarrollo de aplicaciones orientadas a la seguridad personal ha dado lugar a diversas soluciones destinadas a facilitar la solicitud de ayuda y reducir los tiempos de respuesta frente a situaciones de emergencia. Estas



plataformas incorporan mecanismos como activación mediante botones, comandos de voz, ubicación geográfica, redes de contactos de confianza y, más recientemente, técnicas de inteligencia artificial para automatizar la detección de situaciones de riesgo.

Para analizar el contexto tecnológico en el que se desarrolla la mejora de LivIApp, se consideran soluciones que presentan funcionalidades relacionadas con la detección de emergencias, interacción mediante voz, monitoreo de situaciones de riesgo y utilización de inteligencia artificial. El análisis se centra particularmente en los mecanismos empleados para identificar una emergencia y en su capacidad para interpretar información proveniente del usuario.

2.2.1 UrSafe

UrSafe es una aplicación orientada a la seguridad personal que incorpora mecanismos de activación mediante comandos de voz, permitiendo al usuario solicitar asistencia sin necesidad de interactuar directamente con el dispositivo. La plataforma también contempla funcionalidades relacionadas con la ubicación del usuario y la comunicación con contactos de emergencia [9].

Desde la perspectiva de la detección mediante voz, UrSafe representa una aproximación relevante debido a que permite utilizar el lenguaje hablado como mecanismo de interacción en situaciones donde una activación manual puede resultar poco conveniente. Sin embargo, este tipo de mecanismo depende de comandos o expresiones previamente reconocibles por el sistema, lo que limita su capacidad para interpretar situaciones que no sean expresadas mediante patrones esperados.

Esta característica resulta especialmente relevante para LivIApp, debido a que uno de los objetivos de la mejora propuesta consiste precisamente en superar la dependencia de términos estáticos y permitir la interpretación de situaciones de emergencia a partir del contexto de una conversación.

2.2.2 SentiAlert

SentiAlert constituye una solución especialmente relacionada con el problema abordado por LivIApp, debido a que utiliza inteligencia artificial para analizar información de audio con el propósito de identificar situaciones potencialmente peligrosas [10]. A diferencia de los mecanismos basados exclusivamente en una activación manual, este enfoque busca automatizar parte del proceso de detección mediante el análisis del contenido capturado.

La utilización de audio como fuente de información permite abordar situaciones en las que el usuario no necesariamente puede interactuar directamente con su dispositivo. Sin embargo, el análisis automático de conversaciones introduce desafíos asociados a la privacidad, la precisión de la detección y la correcta interpretación del contexto en que ocurre una situación determinada.

Desde la perspectiva de LivIApp, SentiAlert constituye un antecedente relevante al demostrar la viabilidad de utilizar inteligencia artificial para la detección automática de situaciones de riesgo a partir de audio. No obstante, la propuesta de mejora de LivIApp busca avanzar desde la detección de indicadores de riesgo hacia una arquitectura capaz de interpretar y clasificar el contexto de una emergencia, permitiendo ampliar la plataforma hacia múltiples escenarios.

2.2.3 Durcal

Durcal es una plataforma orientada principalmente al monitoreo y seguridad de personas, incorporando funcionalidades como localización, seguimiento y mecanismos de asistencia [11]. Su propuesta se basa en

mantener disponible información relevante sobre la ubicación del usuario y facilitar la comunicación con una red de apoyo ante situaciones que requieran asistencia.

Este tipo de solución demuestra la importancia de combinar la detección de una situación de riesgo con información contextual que permita determinar dónde se encuentra la persona y facilitar una respuesta posterior. Sin embargo, su enfoque se encuentra principalmente orientado al monitoreo y localización, por lo que no constituye una solución centrada en la comprensión semántica de una conversación para determinar automáticamente el tipo de emergencia.

Para LivIApp, este aspecto evidencia que la detección de una emergencia no debe considerarse como un proceso aislado, sino como parte de un flujo que permita recopilar y transmitir información relevante para apoyar la respuesta posterior.

2.2.4 SafeUp

SafeUp corresponde a una aplicación orientada a la seguridad personal que utiliza una red de usuarios y contactos para facilitar la solicitud de asistencia ante situaciones de riesgo. Su funcionamiento se basa principalmente en mecanismos de interacción y comunicación que permiten conectar a una persona que requiere ayuda con una red de apoyo [12].

Este enfoque destaca la importancia de contar con mecanismos que reduzcan el tiempo necesario para establecer contacto con personas capaces de prestar asistencia. Sin embargo, al igual que otras soluciones orientadas principalmente a la activación por parte del usuario, presenta una dependencia de la capacidad de este para identificar la situación y ejecutar el mecanismo correspondiente.

En el contexto de LivIApp, esta limitación resulta relevante debido a que una situación de emergencia puede impedir o dificultar una interacción manual. Por ello, la incorporación de mecanismos de detección automática constituye una oportunidad para reducir la dependencia de la acción consciente del usuario.

2.2.5 Comparación de soluciones

Las soluciones analizadas presentan diferentes aproximaciones al problema de la seguridad personal. Algunas priorizan la activación manual y la comunicación con redes de apoyo, mientras que otras incorporan mecanismos de voz o inteligencia artificial para automatizar parte del proceso de detección. Sin embargo, se observa que las capacidades de comprensión contextual y clasificación automática del tipo de emergencia no constituyen una característica transversal en las soluciones analizadas.

Solución	Voz/Audio	Detección automática	Comprensión Contextual	Multi escenario	Red de apoyo
UrSafe	✓	Parcial	Limitada	✓	✓
SentiAlert	✓	✓	Limitada	Parcial	Parcial
Durcal	-	Parcial	-	✓	✓
SafeUp	-	-	-	✓	✓
LivIApp actual	✓	✓	Limitada	✗	✓
LivIApp propuesta	✓	✓	✓	✓	✓

Tabla 1: Comparativa entre otras aplicaciones y LivIApp. Fuente: Elaboración propia a partir de [9]-[12]



3 Marco Teórico

3.1 Fundamentos Conceptuales

3.1.1 Reconocimiento Automático del Habla (ASR)

El reconocimiento automático del habla (ASR) corresponde al conjunto de técnicas utilizadas para transformar una señal de audio que contiene lenguaje hablado en una representación textual [13]. Este proceso permite incorporar la voz como fuente de información para sistemas informáticos, facilitando posteriormente el análisis del contenido mediante técnicas de procesamiento de lenguaje natural.

En sistemas de detección de emergencias basados en voz, el reconocimiento automático del habla constituye una etapa de entrada, ya que permite transformar el contenido hablado por el usuario en texto que puede ser posteriormente analizado para determinar si existe una situación de riesgo. De esta forma, el proceso de detección puede separarse en dos etapas: una primera etapa orientada a la transcripción del contenido acústico y una segunda encargada de su interpretación semántica.

En la propuesta desarrollada para LivIApp, esta función es realizada mediante Whisper, utilizado para obtener una transcripción textual a partir del audio capturado [1]. La separación entre transcripción e interpretación permite que el contenido generado por el sistema de reconocimiento de voz sea utilizado como entrada para el modelo de lenguaje encargado de determinar la existencia y características de una posible emergencia.

Esta separación resulta relevante desde el punto de vista arquitectónico, debido a que permite desacoplar la adquisición y transformación del audio del proceso posterior de análisis. El flujo utilizado por LivIApp puede representarse conceptualmente como:

Audio capturado -> ASR -> Texto Transcrito -> Análisis Semántico -> Decisión de Emergencia

De esta manera, el reconocimiento automático del habla no constituye por sí mismo el mecanismo de detección de emergencias, sino que proporciona la representación textual necesaria para que una etapa posterior pueda interpretar el contenido de la comunicación.

3.1.2 Procesamiento de Lenguaje Natural

El procesamiento de lenguaje natural (NLP) corresponde al área de la inteligencia artificial orientada al procesamiento y análisis de información expresada mediante lenguaje humano. Su utilización permite que los sistemas informáticos trabajen con información textual de una manera que trasciende la identificación directa de términos, considerando aspectos como el significado, el contexto y la intención expresada en una comunicación [14].

En el contexto de la detección de emergencias, el procesamiento de lenguaje natural permite analizar la transcripción obtenida a partir del audio y determinar si el contenido corresponde a una situación de riesgo. Esto resulta especialmente relevante frente a los enfoques basados exclusivamente en palabras clave, ya que una misma situación puede ser expresada mediante distintas palabras o estructuras lingüísticas. Por lo tanto, el análisis debe considerar el contenido de la comunicación en su conjunto y no únicamente la presencia de términos específicos.

Dentro de LivIApp, el procesamiento de lenguaje natural constituye la etapa posterior al reconocimiento automático del habla. La transcripción generada a partir del audio es entregada a un modelo de lenguaje junto con instrucciones que establecen las categorías de emergencia que deben ser consideradas y los criterios para

determinar si corresponde activar una alerta. De esta manera, el procesamiento puede representarse como una cadena de transformación:

Audio → Transcripción → Interpretación Semántica → Clasificación → Decisión de Activación

En esta arquitectura, el NLP cumple principalmente la función de interpretar la información textual y extraer de ella los elementos relevantes para la toma de decisiones. El resultado de este proceso no corresponde únicamente a una clasificación, sino que incorpora información adicional sobre la categoría y gravedad de la situación, además de una justificación y un resumen destinado a la persona de confianza.

Esta capacidad de transformar lenguaje natural en información que puede ser utilizada por los componentes posteriores del sistema constituye la base para la incorporación de modelos de lenguaje de gran escala, abordados en la siguiente sección.

3.1.3 Modelos de Lenguaje de Gran Escala

Los modelos de lenguaje de gran escala (LLM) son modelos de inteligencia artificial entrenados sobre grandes cantidades de datos textuales, capaces de procesar y generar lenguaje natural. Su capacidad para considerar relaciones entre distintos elementos de una entrada permite abordar tareas que requieren interpretar información más allá de la coincidencia directa de palabras o patrones predefinidos [15].

Una característica relevante de los LLM es su capacidad para realizar distintas tareas mediante instrucciones en lenguaje natural. Entre ellas se encuentran la clasificación de textos, extracción de información, generación de resúmenes y análisis semántico. Esta flexibilidad permite emplear un mismo modelo para abordar diferentes tareas relacionadas con el procesamiento del lenguaje sin requerir necesariamente un modelo especializado para cada una de ellas [15].

En sistemas de detección de emergencias, estas capacidades pueden utilizarse para analizar transcripciones de lenguaje natural y obtener información relevante sobre el contenido de una comunicación. Esto permite abordar escenarios donde la identificación de una situación no depende de la presencia de una palabra específica, sino de la relación entre diferentes elementos presentes en el mensaje y de su contexto.

En este sentido, los LLM constituyen una alternativa a los enfoques basados exclusivamente en reglas o patrones predefinidos, proporcionando capacidades de análisis semántico que pueden ser utilizadas para tareas de clasificación e interpretación de lenguaje natural.

4 Propuesta de Solución

La propuesta de mejora para LivIApp consiste en evolucionar el mecanismo de detección basado en palabras clave hacia una arquitectura capaz de analizar el contenido de comunicaciones de voz y determinar, mediante modelos de lenguaje, si éstas corresponden a una situación de emergencia. La solución busca ampliar el alcance de la plataforma desde el escenario inicial de violencia intrafamiliar hacia un conjunto de situaciones de riesgo, manteniendo la capacidad de generar una alerta para una red de personas de confianza.

Para ello, se incorpora un flujo de procesamiento compuesto por una etapa de captura de audio, una etapa de reconocimiento automático del habla y una etapa de análisis semántico mediante un modelo de lenguaje de gran escala. La información obtenida por cada etapa es transformada progresivamente hasta generar una decisión estructurada que puede ser utilizada por la aplicación para determinar las acciones posteriores.

La arquitectura propuesta busca mantener una separación entre la adquisición de información, su interpretación y la lógica operacional de la aplicación. De esta manera, el modelo de lenguaje se utiliza como componente



especializado en el análisis del contenido de la comunicación, mientras que LivIApp mantiene el control sobre el procesamiento de la respuesta y los mecanismos asociados a la gestión de una emergencia.

4.1 Arquitectura y Diseño

La arquitectura organiza el proceso de detección en distintas etapas, permitiendo transformar una señal de audio en información estructurada que pueda ser utilizada por LivIApp. El flujo general considera la captura del audio, su transcripción mediante un sistema de reconocimiento automático del habla, el análisis de la transcripción mediante un modelo de lenguaje y el procesamiento posterior de la decisión obtenida.

El flujo puede representarse como:

Captura de audio -> Transcripción -> Análisis mediante LLM -> Respuesta estructurada -> Decisión de LivIApp

La separación de estas etapas permite desacoplar las responsabilidades de cada componente. El sistema de reconocimiento automático del habla se encarga exclusivamente de transformar el contenido hablado en texto, mientras que el modelo de lenguaje recibe dicha información y determina sus características relevantes para la detección de emergencias. Finalmente, la aplicación utiliza la respuesta estructurada para continuar con la lógica correspondiente.

4.1.1 Captura y procesamiento del audio

La primera etapa corresponde a la obtención del contenido de voz que será analizado. El audio capturado constituye la entrada inicial del flujo y es posteriormente enviado al componente encargado del reconocimiento automático del habla.

Esta separación permite que el mecanismo de análisis semántico opere sobre una representación textual de la comunicación, evitando que el modelo de lenguaje deba procesar directamente la señal acústica. Como resultado, las etapas de reconocimiento e interpretación mantienen responsabilidades independientes dentro de la arquitectura.

4.1.2 Transcripción mediante Whisper

Una vez obtenido el audio, este es procesado mediante Whisper, utilizado como componente de reconocimiento automático del habla [1]. Su función dentro de la arquitectura consiste en transformar el contenido hablado en una transcripción textual que pueda ser procesada por la siguiente etapa.

Por ejemplo, una entrada de audio correspondiente a una conversación es transformada en una representación textual como:

“Hoy es un lindo día.”

La transcripción constituye, por tanto, la interfaz entre el procesamiento de audio y el análisis semántico. El texto generado es enviado posteriormente al modelo de lenguaje junto con las instrucciones definidas para el proceso de clasificación.

4.1.3 Análisis y clasificación mediante LLM

La transcripción obtenida es procesada mediante un modelo de lenguaje de gran escala, configurado para actuar como motor de análisis de emergencias de LivIApp. En esta etapa, el modelo recibe tanto las instrucciones que definen el comportamiento esperado como la transcripción correspondiente al audio analizado.

El modelo debe determinar si existe una situación de emergencia explícita o implícita y, en caso afirmativo, clasificarla de acuerdo con las categorías definidas para la plataforma. Además, debe proporcionar información adicional que permita caracterizar la situación y facilitar el procesamiento posterior por parte de la aplicación.

En la implementación actual se utiliza GPT-4o-mini, accedido mediante la API de OpenAI [16]. La configuración del modelo restringe su respuesta a una estructura JSON previamente definida, permitiendo que la salida pueda ser procesada de manera determinista por los componentes posteriores de la aplicación.

La respuesta contempla los siguientes atributos:

- **shouldTrigger**: indica si corresponde activar una alerta.
- **categoría**: clasifica la situación como “seguridad”, “médica”, “accidentes” o “ninguna”.
- **gravedad**: establece un nivel “bajo”, “medio” o “alto”.
- **razonamiento**: proporciona una explicación breve de la clasificación.
- **resumenVigilante**: genera un resumen breve destinado a la persona de confianza.

Esta estructura permite separar la interpretación realizada por el modelo de la lógica operacional de LivIApp. El modelo proporciona una clasificación estructurada, mientras que la aplicación puede utilizar dichos valores para determinar las acciones posteriores.

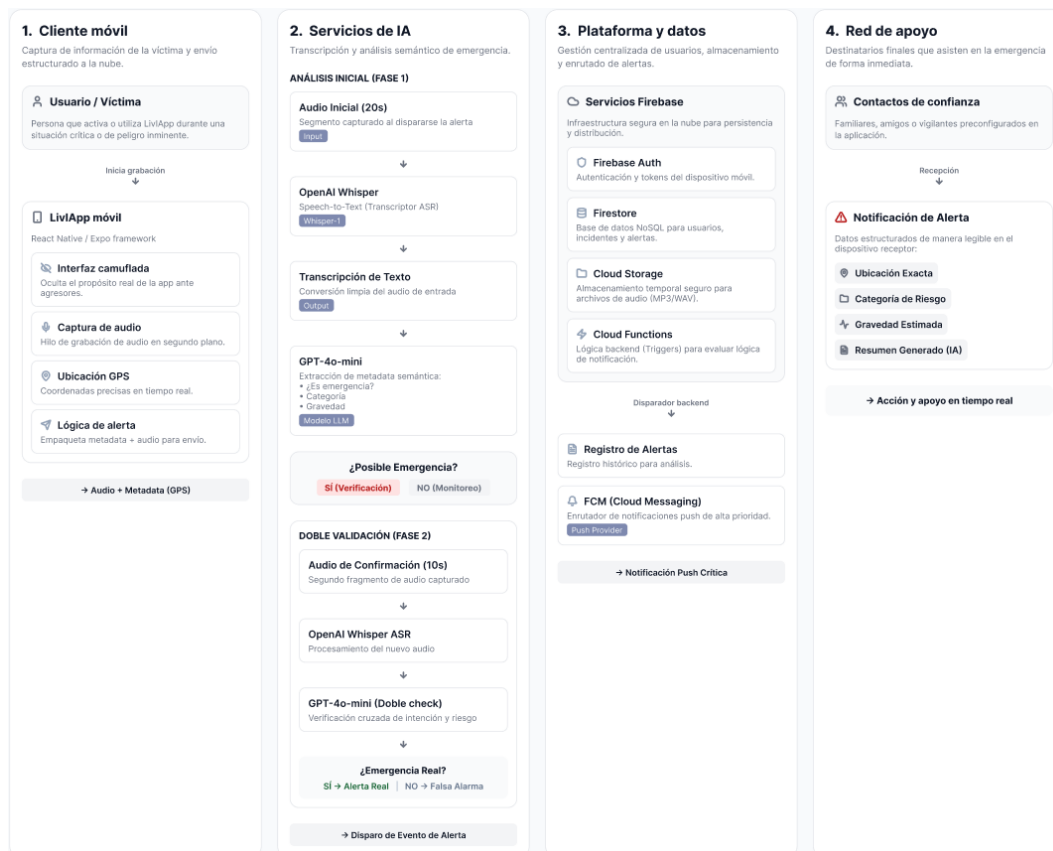


Figura 1: Arquitectura de software propuesta para la detección, verificación y notificación de emergencias en LivIApp.

Fuente: Elaboración propia.

4.2 Integración y Capacidades Técnicas

La solución desarrollada integra los componentes definidos en la arquitectura propuesta en un flujo de procesamiento orientado a la detección y clasificación de situaciones de emergencia mediante información de audio. El proceso combina la captura de segmentos de audio, su transcripción mediante un modelo especializado de reconocimiento de voz y el posterior análisis semántico mediante un modelo de lenguaje. A partir de este análisis, el sistema determina si existe una posible emergencia y, cuando corresponde, ejecuta una segunda etapa de verificación utilizando información contextual adicional.

La integración de estos componentes se implementó como parte del flujo de procesamiento de LivIApp, manteniendo una separación entre las etapas de captura, transcripción, análisis y generación de la respuesta. Esta separación permite que cada componente pueda ser evaluado individualmente y facilita la incorporación de modificaciones en las distintas etapas del sistema.

La arquitectura operativa de LivIApp considera un segmento inicial de 20 segundos y una ventana adicional de 10 segundos para la doble verificación. No obstante, para la batería de pruebas controladas se utilizó un módulo experimental con segmentos de 10 segundos en ambas etapas, con el fin de facilitar la ejecución repetida de los casos de validación.

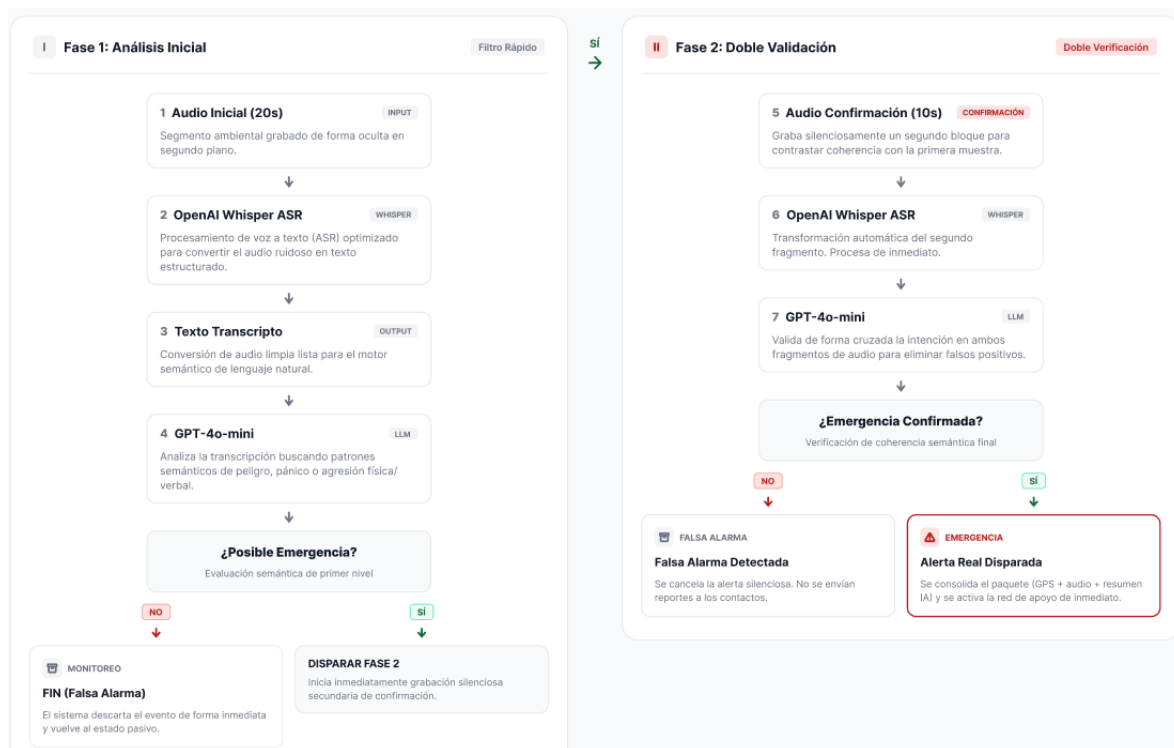


Figura 2: Flujo de procesamiento para la detección y verificación de emergencias en LivIApp. Fuente: Elaboración propia.

El flujo comienza con la captura de un segmento de audio desde el dispositivo móvil, cuyo contenido es almacenado temporalmente y enviado al servicio de transcripción. La transcripción obtenida constituye la entrada para el primer análisis realizado por el modelo de lenguaje. Este análisis determina si existe una posible situación de emergencia y, en caso contrario, el procesamiento finaliza sin generar una alerta.

Cuando el modelo identifica una posible emergencia, LivIApp captura un segmento adicional de audio correspondiente a los segundos posteriores al análisis inicial. Este segundo segmento es procesado mediante el mismo mecanismo de transcripción y posteriormente analizado en conjunto con el contenido inicial. Esta etapa tiene como propósito incorporar información contextual adicional que permita confirmar la existencia de una emergencia o descartar situaciones que inicialmente pudieran haber sido interpretadas como riesgosas.

Finalmente, cuando la segunda etapa confirma la existencia de una emergencia, el sistema genera una respuesta estructurada que contiene la información necesaria para continuar el procesamiento dentro de LivIApp y apoyar la notificación al vigilante correspondiente.

4.2.1 Captura y procesamiento del audio

La primera etapa de la solución corresponde a la captura del audio desde el dispositivo móvil. Para ello, LivIApp utiliza un componente de grabación implementado en React Native [17], permitiendo obtener segmentos de audio que posteriormente son procesados por los servicios de inteligencia artificial.

Los archivos de audio son almacenados temporalmente en formato .m4a utilizando codificación AAC. Una vez finalizada la captura, el archivo se entrega al componente encargado de realizar la transcripción y, posteriormente, es eliminado del almacenamiento temporal. De esta forma, el archivo utilizado durante el procesamiento no permanece innecesariamente en el dispositivo.

El sistema trabaja de manera segmentada, permitiendo analizar el contenido progresivamente en lugar de requerir la grabación completa de una conversación. Este mecanismo resulta relevante para un escenario de detección en tiempo cercano al real, ya que permite iniciar el procesamiento mientras la situación continúa desarrollándose.

En la implementación de la doble verificación, cuando el primer análisis determina que existe una posible emergencia, se captura un segundo segmento de audio de 10 segundos. Este segmento se utiliza como contexto adicional para la etapa de confirmación.

4.2.2 Transcripción mediante Whisper

Una vez obtenido el segmento de audio, este es enviado al servicio de transcripción de OpenAI mediante el modelo Whisper [1]. La implementación especifica el idioma español para el procesamiento de las grabaciones, considerando que LivIApp está orientada inicialmente a usuarios de habla hispana.

El resultado de esta etapa corresponde a una transcripción textual del segmento de audio, la cual constituye la entrada utilizada posteriormente por el modelo de lenguaje. De esta manera, la solución separa el reconocimiento del contenido hablado de su interpretación semántica, permitiendo que cada etapa tenga una responsabilidad específica dentro del flujo.

La separación entre transcripción y análisis semántico permite además evaluar ambas etapas de manera independiente durante la validación de la solución, particularmente en aspectos relacionados con la calidad de la transcripción y los tiempos de procesamiento.

4.2.3 Análisis semántico mediante LLM

La interpretación de la transcripción se realiza mediante el modelo de lenguaje GPT-4o-mini [16], utilizando un prompt diseñado específicamente para el contexto de detección de emergencias de LivIApp.



El modelo recibe la transcripción obtenida desde Whisper y debe determinar si el contenido corresponde a una situación de emergencia explícita o implícita. Para ello, el prompt establece cuatro categorías posibles: seguridad, médica, accidentes y ninguna. Además, solicita determinar un nivel de gravedad entre baja, media y alta.

La respuesta del modelo se restringe a una estructura JSON, permitiendo que el resultado pueda ser procesado programáticamente por la aplicación. La estructura considera los campos “shouldTrigger”, “categoria”, “gravedad”, “razonamiento” y “resumenVigilante”. El campo “shouldTrigger” permite determinar si el flujo debe continuar hacia la etapa de verificación, mientras que los demás campos permiten estructurar la información obtenida del análisis.

Para reducir la variabilidad de las respuestas generadas, la implementación utiliza una temperatura de 0,1 y el mecanismo de respuesta estructurada proporcionado por la API. De esta forma, el resultado del modelo puede ser interpretado directamente por la aplicación sin depender de formatos textuales variables.

4.2.4 Mecanismo de doble verificación

Una de las capacidades incorporadas a la solución corresponde a un mecanismo de doble verificación orientado principalmente a la mitigación de falsos positivos. Este mecanismo se ejecuta cuando el primer análisis identifica una posible situación de emergencia.

Ante una sospecha inicial, LivIApp captura un segundo segmento de audio de 10 segundos. Este nuevo contenido es transcrito mediante Whisper y posteriormente enviado al modelo de lenguaje junto con la transcripción obtenida durante el análisis inicial.

A diferencia del primer análisis, la segunda etapa no busca únicamente identificar una categoría de emergencia, sino determinar si la situación detectada inicialmente corresponde efectivamente a una emergencia real. Para ello, el modelo recibe ambos segmentos como contexto y genera una respuesta estructurada mediante los campos “isRealEmergency”, razonamiento y “gravedadFinal”.

Este mecanismo permite que información inicialmente ambigua pueda ser contrastada con el contexto inmediatamente posterior. De esta manera, situaciones como conversaciones cotidianas, referencias a emergencias ocurridas en el pasado o contenido proveniente de medios audiovisuales pueden ser descartadas antes de generar una alerta definitiva.

Asimismo, la segunda etapa puede aportar información adicional para la evaluación de gravedad. En consecuencia, el mecanismo no se limita exclusivamente a reducir falsos positivos, sino que también permite refinar la interpretación de una situación cuando el segundo segmento proporciona antecedentes relevantes.

4.2.5 Generación de respuesta y trazabilidad

Una vez completado el análisis, la información obtenida se estructura para ser utilizada por LivIApp. En caso de confirmarse una emergencia, el sistema dispone de una respuesta que identifica la categoría y gravedad estimadas, junto con un razonamiento breve y un resumen destinado al vigilante.

El campo “resumenVigilante” permite transformar el resultado del análisis en información concisa que pueda ser utilizada posteriormente en el flujo de notificación de LivIApp. Esto evita que el receptor de la alerta deba interpretar directamente la salida generada por el modelo de lenguaje.

Además, la estructura de respuesta permite mantener trazabilidad sobre las decisiones tomadas durante el procesamiento. El registro de elementos como la categoría, gravedad, resultado de la verificación y tiempos

asociados al procesamiento permite analizar posteriormente el comportamiento del sistema y facilita la evaluación de su funcionamiento.

La incorporación de esta información resulta especialmente relevante para la etapa de validación, donde se requiere analizar no solo si el sistema genera una alerta correctamente, sino también cómo se comportan las distintas etapas que intervienen en la decisión final.

5 Metodología de Validación y Resultados

5.1 Diseño de la Validación

5.1.1 Objetivo de la Validación

La validación tiene como propósito evaluar el funcionamiento de la arquitectura desarrollada para la detección y clasificación de situaciones de emergencia mediante lenguaje natural, considerando tanto el comportamiento del análisis inicial como el mecanismo de doble verificación incorporado para reducir falsos positivos.

En particular, se busca determinar si la solución es capaz de distinguir entre situaciones de emergencia y conversaciones que no representan un riesgo, clasificar correctamente el tipo de emergencia identificada y determinar su nivel de gravedad. Adicionalmente, se evalúa el aporte de la segunda etapa de análisis contextual y el tiempo requerido para procesar los segmentos de audio.

La validación también busca contrastar el comportamiento de la implementación con los objetivos definidos para el proyecto, particularmente respecto de la capacidad de detección multi escenario, el procesamiento de los casos de prueba y la latencia del flujo de análisis.

5.1.2 Ambiente y Condiciones de Prueba

Las pruebas fueron realizadas en un ambiente controlado, utilizando la implementación funcional de LivIApp y un conjunto de situaciones simuladas diseñadas específicamente para representar diferentes escenarios de emergencia y situaciones que podrían generar una detección incorrecta.

Para cada caso se utilizaron segmentos de audio correspondientes a conversaciones previamente definidas. El contenido fue procesado mediante el flujo implementado en LivIApp, compuesto por la captura del audio, transcripción mediante Whisper y análisis semántico mediante GPT-4o-mini. Cuando el análisis inicial identificó una posible emergencia, se ejecutó la etapa de doble verificación mediante un segundo segmento de audio y un nuevo análisis contextual.

Las pruebas fueron realizadas bajo las mismas condiciones generales de funcionamiento del sistema, registrando para cada caso los resultados obtenidos en las distintas etapas del procesamiento. En la segunda batería de pruebas también se registraron los tiempos asociados al análisis de cada segmento, con el objetivo de evaluar la latencia del flujo.

Es importante señalar que los casos utilizados corresponden a escenarios simulados y controlados, por lo que los resultados obtenidos permiten evaluar el comportamiento de la implementación bajo las condiciones establecidas, pero no constituyen una estimación directa de su desempeño ante emergencias reales.

5.1.3 Diseño de los Casos de Prueba

La validación se estructuró en **dos baterías** de pruebas complementarias, alcanzando un total de 40 casos.



La **primera batería** estuvo compuesta por 30 casos, distribuidos entre las tres categorías de emergencia contempladas por la solución: seguridad, médica y accidentes. Para cada categoría se definieron cinco casos correspondientes a situaciones de emergencia y cinco casos correspondientes a situaciones de no emergencia. Los casos negativos fueron diseñados para representar situaciones que contienen expresiones potencialmente asociadas a una emergencia, pero que corresponden a conversaciones cotidianas, hechos ocurridos en el pasado o contenido proveniente de medios audiovisuales. Los resultados fueron registrados en la planilla de validación correspondiente.

La **segunda batería** estuvo compuesta por 10 casos adicionales, diseñados específicamente para evaluar el comportamiento del mecanismo de doble verificación. Estos casos consideraron situaciones donde el primer segmento contenía información suficiente para generar una decisión, casos ambiguos que requerían contexto adicional, situaciones donde la gravedad podía modificarse después de incorporar nueva información y casos en los que el segundo segmento permitía descartar una sospecha inicial. Para cada prueba se registró el resultado del primer análisis, la necesidad de utilizar un segundo segmento, la decisión final, la variación de gravedad y los tiempos de procesamiento.

Esta segunda batería permite complementar la evaluación de exactitud realizada mediante los primeros 30 casos, incorporando una evaluación específica del comportamiento interno del flujo de decisión.

Cabe destacar que los 40 casos considerados en esta validación fueron definidos y ejecutados una vez finalizada la configuración del prompt y del flujo de procesamiento de la solución, y fueron distintos de los casos utilizados durante la etapa de desarrollo y ajuste del sistema. Estos escenarios no fueron utilizados posteriormente para modificar las instrucciones del modelo ni reajustar su lógica de decisión. De esta manera, el conjunto de validación se mantuvo separado de la etapa de configuración del mecanismo de análisis. No obstante, dado que corresponde a escenarios simulados y ejecutados bajo condiciones controladas, sus resultados se interpretan como una validación funcional de la implementación y no como una estimación generalizable del desempeño en condiciones reales.

5.1.4 Criterios y Métricas de Evaluación

Para determinar la correcta ejecución de los casos de prueba se establecieron criterios de evaluación asociados a la decisión final del sistema. En los casos definidos como emergencias, se considera correcta una respuesta cuando el sistema confirma la existencia de una emergencia. En los casos negativos, se considera correcta cuando el sistema descarta la situación y evita la generación de una alerta.

Para evaluar la capacidad de clasificación multi escenario, en los casos positivos se comparó además la categoría esperada definida para cada escenario con la categoría entregada por el modelo, considerando las categorías seguridad, médica y accidentes.

A partir de estas decisiones se consideran las siguientes métricas:

- Verdaderos positivos (VP): casos de emergencia correctamente detectados.
- Verdaderos negativos (VN): casos sin emergencia correctamente descartados.
- Falsos positivos (FP): casos sin emergencia en los que el sistema determina incorrectamente que existe una emergencia.
- Falsos negativos (FN): casos de emergencia que el sistema no logra detectar una emergencia.

A partir de estos valores se calcularon **accuracy**, **precision**, **recall** y **F1-score**, permitiendo evaluar el desempeño general de la detección [18].

Además, se analizó descriptivamente el comportamiento de la clasificación y de la gravedad obtenida, considerando su correspondencia con las características definidas para cada escenario.

Finalmente, para evaluar el desempeño temporal de la solución se considera la latencia de procesamiento posterior a la captura de audio. La medición comienza una vez finalizado cada segmento de grabación y concluye al obtener el resultado correspondiente del análisis, por lo que no incluye la duración de las ventanas de captura. En los casos que ejecutan la doble verificación, se considera además el procesamiento acumulado de ambas etapas.

5.1.5 Procedimiento Experimental

Cada caso de prueba fue ejecutado siguiendo el flujo definido por la arquitectura de la solución. Inicialmente se capturó un segmento de audio y se obtuvo su correspondiente transcripción mediante Whisper. Posteriormente, la transcripción fue enviada al modelo de lenguaje para determinar si existía una posible situación de emergencia y, cuando correspondía, identificar su categoría y gravedad.

Cuando el resultado inicial indicó una posible emergencia, el sistema procedió a capturar y procesar un segundo segmento de audio. Este contenido fue transcrito y analizado conjuntamente con la información inicial mediante el mecanismo de doble verificación. El resultado de esta etapa determinó si la situación correspondía a una emergencia real o a una falsa alarma.

Para cada ejecución se registraron los resultados obtenidos y se compararon con el resultado esperado definido previamente para el caso. En la segunda batería se incorporó además el registro de los tiempos de procesamiento de cada etapa.

Las capturas de las ejecuciones realizadas durante la validación se presentan en los anexos como evidencia complementaria del proceso experimental.

5.2 Análisis de Resultados

5.2.1 Desempeño general de la detección

La primera batería de validación estuvo compuesta por 30 casos de prueba, distribuidos entre situaciones de seguridad, médicas y accidentes, además de casos negativos diseñados para representar situaciones que podían contener expresiones asociadas a emergencias sin corresponder a un riesgo real. En esta batería, los 30 casos obtuvieron una decisión final coincidente con el resultado esperado.

De los 30 casos evaluados, 15 corresponden a situaciones de emergencia y 15 a situaciones de no emergencia. Los casos positivos fueron correctamente identificados, mientras que los casos negativos fueron clasificados como falsas alarmas, evitando la generación de una alerta.

Resultado	Casos
Casos evaluados	30
Emergencias esperadas	15
Casos sin emergencia	15
Verdaderos positivos	15
Verdaderos negativos	15
Falsos positivos	0
Falsos negativos	0

Tabla 2: Resultados generales de la primera batería de pruebas. Fuente: Elaboración propia

A partir de estos resultados, el sistema obtuvo una exactitud de:



$$Accuracy = \frac{VP + VN}{VP + VN + FP + FN}$$
$$Accuracy = 15 + 1515 + 15 + 0 + 0 = 1 = 100\%$$

Dentro de esta primera batería de escenarios simulados, la decisión final alcanzó una accuracy, precision, recall y F1-score de 100 %. Estos valores deben interpretarse considerando las condiciones de la validación. La muestra corresponde a casos diseñados y ejecutados en un ambiente controlado, por lo que no permite afirmar que el sistema mantendría el mismo desempeño frente a la diversidad y variabilidad presentes en situaciones reales.

5.2.2 Desempeño según Categoría de Emergencia

Además del resultado general, se analizó el comportamiento de la solución según el tipo de emergencia representado en los casos positivos. La batería consideró situaciones de seguridad, médicas y accidentes, incluyendo algunos escenarios que presentan características compartidas entre accidentes y emergencias médicas.

En los casos de seguridad se evaluaron situaciones como intrusiones, amenazas y seguimiento. En los casos médicos se consideraron, entre otros, dificultades respiratorias, pérdida de conciencia y síntomas físicos. Finalmente, los casos de accidentes incluyeron situaciones como incendios, fugas de gas, caídas y quemaduras.

Con el objetivo de demostrar directamente la capacidad de clasificación multi escenario, la Tabla 3 presenta la comparación entre la categoría esperada y la categoría obtenida por el sistema para los 15 casos positivos evaluados. La comparación se estructura como una matriz de confusión para las categorías seguridad, médica y accidentes.

Categoría esperada	Casos evaluados	Seguridad obtenida	Médica obtenida	Accidentes obtenida	Coincidencias
Seguridad	5	5	0	0	5/5
Médica	5	0	5	0	5/5
Accidentes	5	0	0	5	5/5
Total	15	5	5	5	15/15

Tabla 3: Comparación entre categoría esperada y categoría obtenida para los escenarios positivos de validación.

Fuente: Elaboración propia

Como se observa en la Tabla 3, los cinco casos esperados de seguridad fueron clasificados como seguridad, los cinco casos médicos fueron clasificados en la categoría médica y los cinco casos de accidentes como accidentes, obteniéndose 15 de 15 coincidencias dentro del conjunto controlado de validación. Este resultado permitió verificar funcionalmente la capacidad de clasificación multi escenario de la solución para las categorías de seguridad, médica y accidentes. En los casos que presentaban características compartidas entre accidentes y emergencias médicas, se mantuvo como referencia la categoría esperada definida previamente para la validación, reconociendo que este tipo de situaciones puede admitir más de una interpretación. Estos resultados deben considerarse exclusivamente dentro de las condiciones controladas de evaluación y no como una estimación general del desempeño del sistema en escenarios reales.

5.2.3 Evaluación del mecanismo de doble verificación

La segunda batería estuvo compuesta por 10 casos y tuvo como objetivo evaluar específicamente el comportamiento del mecanismo de doble verificación. A diferencia de la primera batería, estos casos fueron seleccionados para observar qué ocurre cuando el primer análisis genera una sospecha que requiere información contextual adicional.

De los 10 casos evaluados, 9 activaron la segunda etapa de procesamiento, mientras que un caso fue descartado directamente durante el análisis inicial. En los nueve casos procesados mediante doble verificación, la decisión final coincidió con el resultado esperado.

Un resultado particularmente relevante corresponde a los casos negativos 34, 36 y 40. En ellos, el análisis inicial identificó una posible situación de emergencia, pero el contenido del segundo segmento permitió determinar que no existía una emergencia real. Por ejemplo, en uno de los casos una expresión inicial de necesidad de ayuda fue posteriormente contextualizada como una situación relacionada con carga de trabajo, mientras que en otro la expresión asociada a un incendio correspondía a una escena de una película.

Indicador	Resultado
Casos evaluados	10
Emergencias reales	6
Casos sin emergencia	4
Casos que utilizaron Audio 2	9
Casos descartados en Audio 1	1
Falsos positivos finales	0
Falsos negativos finales	0

Tabla 4: Resultados de la segunda batería. Fuente: Elaboración propia

El comportamiento observado permite identificar una diferencia entre el resultado del análisis inicial y la decisión final del sistema. En particular, algunos casos negativos generaron una sospecha inicial debido a la presencia de expresiones que, aisladas de su contexto, podrían interpretarse como indicadores de emergencia. La segunda etapa permitió contrastar esta interpretación con información adicional y descartar dichas situaciones.

Por lo tanto, la doble verificación cumplió el propósito para el cual fue incorporada a la arquitectura: reducir el riesgo de falsos positivos mediante el análisis de contexto adicional antes de generar una alerta definitiva.

5.2.4 Impacto de la información contextual en la decisión

Además de permitir descartar falsas alarmas, la segunda etapa produjo cambios en la evaluación de gravedad de algunos casos positivos. En particular, los casos 38 y 39 pasaron desde una gravedad media durante el primer análisis a una gravedad alta después de incorporar el segundo segmento de audio.

Este comportamiento muestra que el segundo segmento no funciona únicamente como un mecanismo binario de confirmación, sino que puede proporcionar información adicional para caracterizar la evolución de una situación.

Por ejemplo, en uno de los casos médicos el primer segmento permitió identificar una posible emergencia, mientras que el segundo informó que la persona comenzaba a sentirse incapaz de mantenerse en pie y próxima a desmayarse. De forma similar, en un caso de accidente, el aumento del dolor posterior permitió al modelo elevar la gravedad estimada.

En consecuencia, la incorporación de contexto temporal inmediato permite complementar la información obtenida en el primer análisis y puede contribuir tanto a la reducción de falsos positivos como al refinamiento de la gravedad asignada.

5.2.5 Análisis de latencia

La segunda batería también permitió medir el tiempo requerido por las etapas de procesamiento. Para cada caso se registró la latencia asociada al primer análisis, al segundo análisis cuando correspondía y al procesamiento total.

Considerando los nueve casos que utilizaron ambas etapas, la latencia promedio del primer análisis fue de 4,12 segundos, mientras que la segunda etapa presentó una latencia promedio de 3,75 segundos.

El procesamiento acumulado de ambas etapas presentó una latencia promedio de aproximadamente 7,87 segundos.

Métrica	Tiempo promedio [s]
Procesamiento posterior al Audio 1	4,12
Procesamiento posterior al Audio 2	3,75
Procesamiento acumulado de ambas etapas	7,87

Tabla 5: Latencia promedio de la segunda batería. Fuente: Elaboración propia

Estos tiempos corresponden exclusivamente al procesamiento ejecutado después de finalizar cada captura de audio y no incluyen los 10 segundos utilizados para cada ventana de grabación en el módulo experimental. Por tanto, los 7,87 segundos representan la latencia acumulada de procesamiento de ambas etapas y no un tiempo extremo a extremo desde el inicio de la captura hasta la generación de la alerta.

6 Conclusiones

6.1 Conclusiones

El trabajo desarrollado permitió diseñar e implementar una arquitectura de software orientada a ampliar las capacidades de detección de emergencias de LivIApp, reemplazando la dependencia exclusiva de palabras clave por un flujo basado en transcripción automática y análisis semántico mediante modelos de lenguaje. La solución integra la captura de audio, transcripción mediante Whisper, clasificación mediante GPT-4o-mini y un mecanismo de doble verificación que utiliza información contextual adicional antes de generar una alerta.

Respecto de los objetivos planteados, se implementó un pipeline capaz de transformar información de voz en una respuesta estructurada utilizable por LivIApp, incorporando la identificación de situaciones de seguridad, emergencias médicas y accidentes, junto con una estimación de gravedad y un resumen destinado a la red de apoyo. Asimismo, el mecanismo de doble verificación permitió incorporar contexto temporal posterior para confirmar situaciones inicialmente sospechosas o descartarlas antes de generar una alerta definitiva.

La validación realizada en ambiente controlado mostró resultados favorables. En la primera batería, los 30 casos evaluados obtuvieron una decisión final coincidente con el resultado esperado, sin falsos positivos ni falsos negativos dentro del conjunto considerado. Adicionalmente, en los escenarios positivos la comparación entre categoría esperada y obtenida permitió evaluar explícitamente la capacidad de clasificación multi escenario, obteniéndose coincidencia en los 15 casos positivos evaluados dentro del conjunto controlado de validación. En la segunda batería, 9 de los 10 casos activaron el mecanismo de doble verificación y todas las decisiones finales coincidieron con los resultados esperados.

Desde el punto de vista temporal, el procesamiento acumulado de las dos etapas de análisis presentó una latencia promedio de aproximadamente 7,87 segundos bajo las condiciones utilizadas en las pruebas, manteniéndose por debajo del umbral de 10 segundos definido para la solución. Esta medición considera el tiempo transcurrido

desde la finalización de cada captura de audio hasta la obtención de su correspondiente resultado, incluyendo en la segunda etapa la activación de la alerta, pero excluyendo la duración de las ventanas de grabación.

En conjunto, los resultados entregan evidencia inicial de que una arquitectura basada en reconocimiento automático del habla y modelos de lenguaje puede ampliar las capacidades de LivIApp hacia un escenario multi escenario, incorporando mecanismos de interpretación contextual que no están presentes en un enfoque basado exclusivamente en términos predefinidos. La principal contribución del trabajo corresponde, por tanto, al diseño e integración de este flujo de procesamiento dentro de LivIApp, junto con la incorporación de un mecanismo de doble verificación y su evaluación funcional mediante escenarios controlados.

6.2 Limitaciones y Oportunidades de Mejora

A pesar de los resultados positivos obtenidos, la validación presenta algunas limitaciones que deben ser consideradas al interpretar los resultados.

En primer lugar, los casos fueron ejecutados en un ambiente controlado, utilizando situaciones previamente diseñadas. Esto permite reproducir las pruebas y comparar los resultados con condiciones esperadas, pero no representa completamente la variabilidad del lenguaje espontáneo, las condiciones acústicas y los factores emocionales presentes durante una emergencia real.

En segundo lugar, el número de casos utilizado permite realizar una primera evaluación funcional de la arquitectura, pero resulta insuficiente para generalizar estadísticamente el comportamiento del sistema a una población amplia de situaciones de emergencia.

Finalmente, los resultados de la segunda batería muestran una oportunidad de mejora en el análisis inicial. Algunos casos que posteriormente fueron descartados como falsas alarmas generaron inicialmente una sospecha y requirieron una segunda llamada al modelo. Aunque el mecanismo de doble verificación permitió corregir estos casos sin generar una alerta final incorrecta, una futura optimización del prompt podría mejorar la precisión de la primera etapa y reducir la cantidad de verificaciones innecesarias.

Esta mejora debe realizarse manteniendo un equilibrio entre la reducción de falsos positivos y la capacidad de detectar situaciones de emergencia expresadas de manera ambigua, evitando que un criterio excesivamente restrictivo incremente los falsos negativos.

Como trabajo futuro, resulta pertinente ampliar la validación utilizando un conjunto mayor y más diverso de escenarios, incorporando variaciones de ruido ambiental, calidad del audio y formas espontáneas de expresión. Asimismo, los resultados obtenidos sugieren la posibilidad de refinar las instrucciones utilizadas durante el primer análisis para reducir verificaciones innecesarias sin incrementar el riesgo de falsos negativos. Finalmente, podrían explorarse mecanismos adicionales de protección de la información y alternativas que reduzcan la dependencia de servicios externos, especialmente considerando la sensibilidad de los datos procesados y la necesidad de mantener disponibilidad en situaciones críticas.

6.3 Agradecimientos.

Quiero agradecer a todas las personas que me acompañaron durante este largo y significativo proceso.

A mi profesor guía, Mauricio Figueroa, por su orientación y por ayudarme a enfocar de mejor manera el desarrollo de esta memoria.

A los docentes que estuvieron presentes durante mi formación, por su tiempo, disposición y enseñanza. Al equipo de LivIApp, por el trabajo compartido y por una experiencia que trascendió lo profesional para convertirse también en amistad.



A mi familia, por su apoyo constante, por confiar en mí y por motivarme a seguir avanzando incluso en los momentos más difíciles. Gracias por el apoyo de los que estuvieron y de los que ya no están.

Y especialmente a mi pareja, quien durante estos años ha sido un pilar fundamental, acompañándome en gran parte de este camino, junto a nuestra Lunita, brindándome apoyo, paciencia y contención cuando más lo necesité.

Referencias

- [1] J. R. Batista, *Learn OpenAI Whisper*, 1st ed. Birmingham, U.K.: Packt Publishing, 2024. [En línea]. Disponible en: https://books.google.cl/books?hl=es&lr=&id=QVoIEQAAQBAJ&oi=fnd&pg=PP1&dq=Whisper+OpenAI&ots=232PZ4sT_2&sig=5clswaBQaxsnhN2FWVZvPNcMbr0&redir_esc=y#v=onepage&q=Whisper%20OpenAI&f=false. [Accedido: Agosto 2026]
- [2] V. Alto, *Inteligencia artificial generativa con modelos de ChatGPT y OpenAI*, 1.^a ed. Madrid, España: Anaya Multimedia, 2023. [En línea]. Disponible en: https://books.google.cl/books?id=nzSgEQAAQBAJ&pg=PP33&dq=LLM+OPENAI&lr=&hl=es&source=gbs_selected_pages&cad=1#v=onepage&q=LLM%20OPENAI&f=false. [Accedido: Agosto 2026]
- [3] S. Fernández, A. Graves and J. Schmidhuber, “An application of recurrent neural networks to discriminative keyword spotting,” in *Artificial Neural Networks—ICANN 2007, Lecture Notes in Computer Science*, vol. 4669, pp. 220–229, 2007, doi: 10.1007/978-3-540-74695-9_23.
- [4] J. A. Moyano Dután and A. I. Roman Alvarado, “Creación de una infraestructura auto adaptativa de monitoreo mediante un middleware de configuración de componentes sensibles al contexto en ambientes de vida asistida,” *Trabajo de Titulación, Universidad de Cuenca, Cuenca, Ecuador*, 2021. [En línea]. Disponible en: <https://dspace.ucuenca.edu.ec/items/a1510d61-a7c0-46f7-965a-36a402d43406>. [Accedido: Agosto 2026].
- [5] P. Brusco, “Keyword spotting en idiomas sin datos de entrenamiento,” *Tesis de Licenciatura, Departamento de Computación, Universidad de Buenos Aires, Buenos Aires, Argentina*, 2014. [En línea]. Disponible en: https://bibliotecadigital.exactas.uba.ar/download/seminario/seminario_nCOM000705_Brusco.pdf. [Accedido: Agosto 2026].
- [6] Y. Kang, Z. Cai, C.-W. Tan, Q. Huang and H. Liu, “Natural language processing (NLP) in management research: A literature review,” *Journal of Management Analytics*, vol. 7, no. 2, pp. 139–172, 2020, doi: 10.1080/23270012.2020.1756939.
- [7] D. L. Paltín Chica, J. D. Mejía Mendieta, M. Orellana and J. L. Zambrano-Martínez, “Modelado semántico de emergencias del ECU 911 con NLP y ontologías,” *Revista Tecnológica - ESPOL*, vol. 37, no. E1, 2025, doi: 10.37815/rte.v37nE1.1351.
- [8] S. Errigo and J. Palacios, “Análisis de requerimientos para una app de alerta de emergencias médicas en la UTP Centro Regional de Chiriquí,” *Memorias de Congresos UTP*, pp. 146–152, 2019. [En línea]. Disponible en: <https://revistas.utp.ac.pa/index.php/memoutp/article/view/2307>. [Accedido: Agosto 2026].
- [9] UrSafe, “UrSafe Personal Safety App.” [En línea]. Disponible en: <https://ursafe.com/>. [Accedido: Agosto 2026].
- [10] Universidad Técnica Federico Santa María, “SentIAAlert,” *Feria de Software*. [En línea]. Disponible en: <https://www.feriadesoftware.cl/proyectos/sentialert-2>. [Accedido: Septiembre 2026].
- [11] Durcal, “Reloj de teleasistencia para ancianos y personas mayores.” [En línea]. Disponible en: <https://www.durcal.com/>. [Accedido: Septiembre 2026].
- [12] A. Martínez Ortega, “Crean SafeUP con ‘guardianas’ para prevenir la violencia contra mujeres,” *La Opinión*, 5 mayo 2021. [En línea]. Disponible en: <https://laopinion.com/2021/05/05/crean-safeup-con-guardianas-para-prevenir-la-violencia-contra-mujeres/>. [Accedido: Septiembre 2026].
- [13] H. Strik and C. Cucchiari, “Modeling pronunciation variation for ASR: A survey of the literature,” *Speech Communication*, vol. 29, nos. 2–4, pp. 225–246, 1999, doi: 10.1016/S0167-6393(99)00038-2.
- [14] Universidad Politécnica de Madrid, “NLP y recuperación de información,” *Guía de aprendizaje, curso 2026–2027*. [En línea]. Disponible en: https://www.upm.es/comun_gauss/publico/guias/2026-27/GA_09BM_95000346_ES_2026-27.pdf. [Accedido: Septiembre 2026].
- [15] J. C. Jiménez Revuelta, “Modelos Grandes de Lenguaje y aplicaciones a la generación automática de texto,” *Trabajo Fin de Máster, Universidad de Sevilla, Sevilla, España*, 2023. [En línea]. Disponible en: <https://idus.us.es/server/api/core/bitstreams/c9feaa9c-5acf-4ba1-b53a-84da6ec092ae/content>. [Accedido: Septiembre 2026].

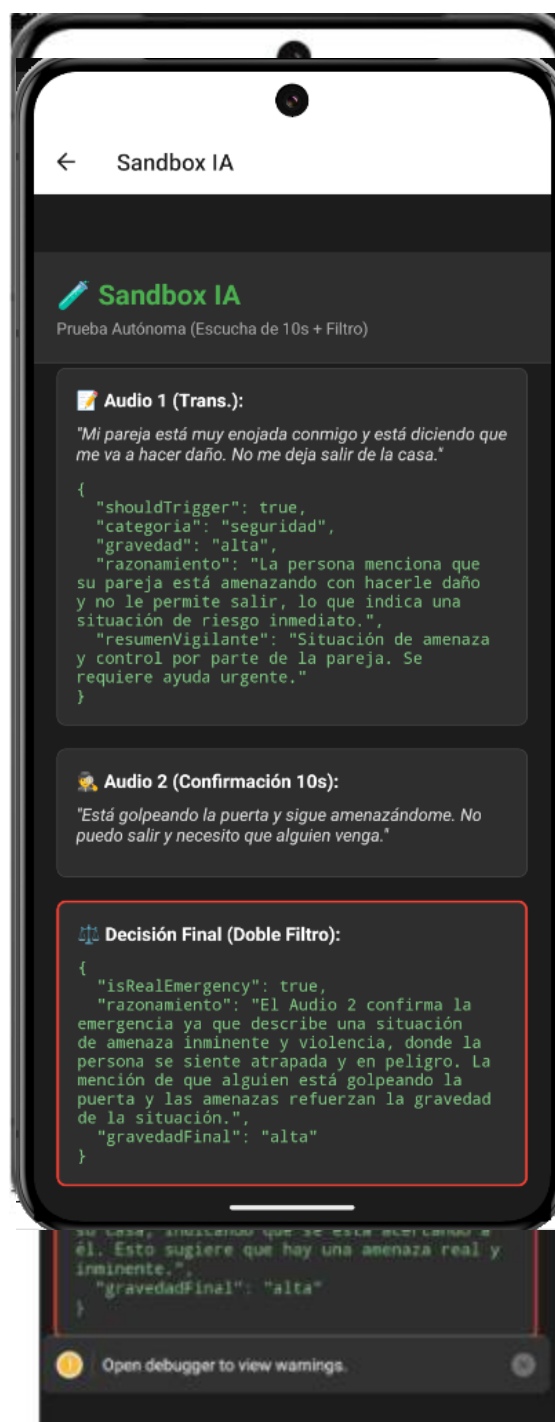


- [16] OpenAI, “GPT-4o mini: advancing cost-efficient intelligence,” 18 julio 2024. [En línea]. Disponible en: <https://openai.com/es-419/index/gpt-4o-mini-advancing-cost-efficient-intelligence/>. [Accedido: Septiembre 2026].
- [17] ProTech Training, “React Native,” curso PT20726. [En línea]. Disponible en: <https://www.protechtraining.com/pdf/PT20726~.pdf>. [Accedido: Septiembre 2026].
- [18] C. Goutte and É. Gaussier, “A probabilistic interpretation of precision, recall and F-score, with implication for evaluation,” in Advances in Information Retrieval: 27th European Conference on IR Research, ECIR 2005, Lecture Notes in Computer Science, vol. 3408, pp. 345–359, 2005, doi: 10.1007/978-3-540-31865-1_25.

Anexos

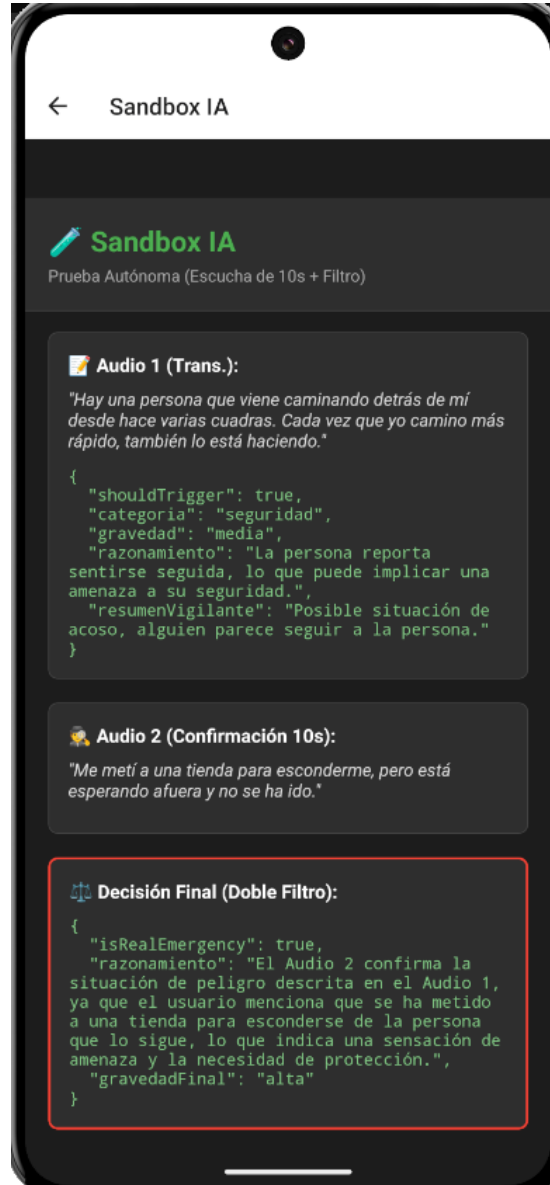
Anexo A. Evidencias de la primera batería de validación

Caso 1:

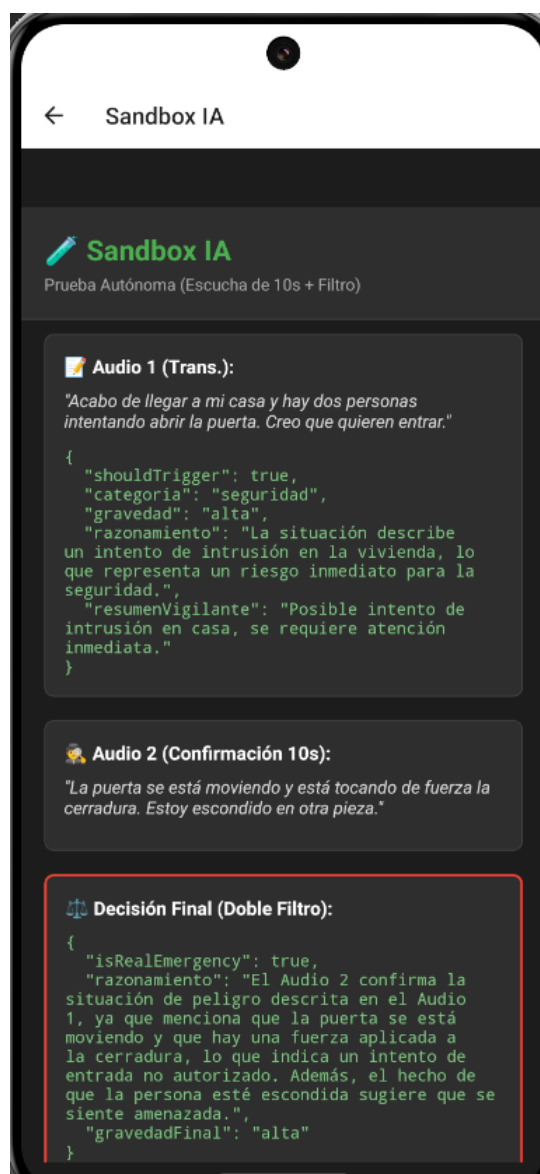


Caso 2:

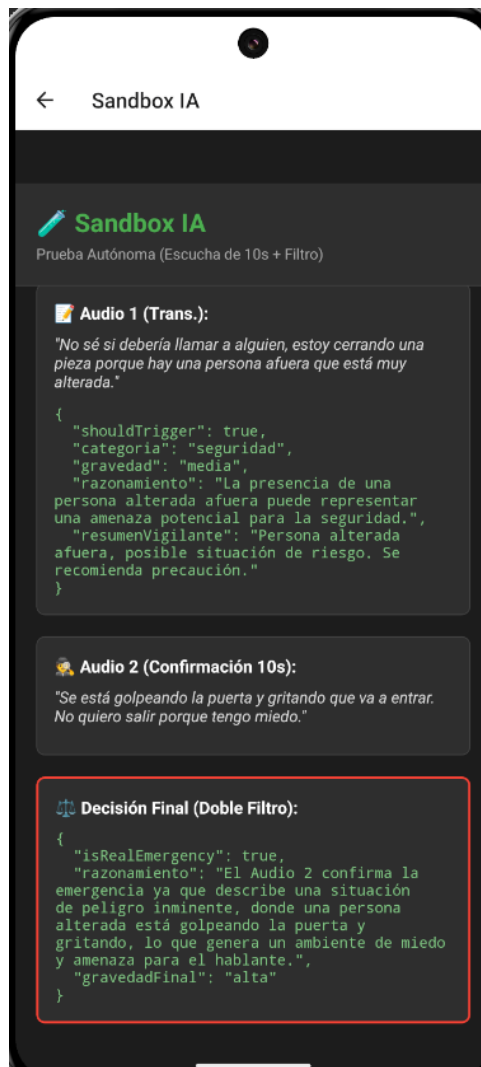
Caso 3:



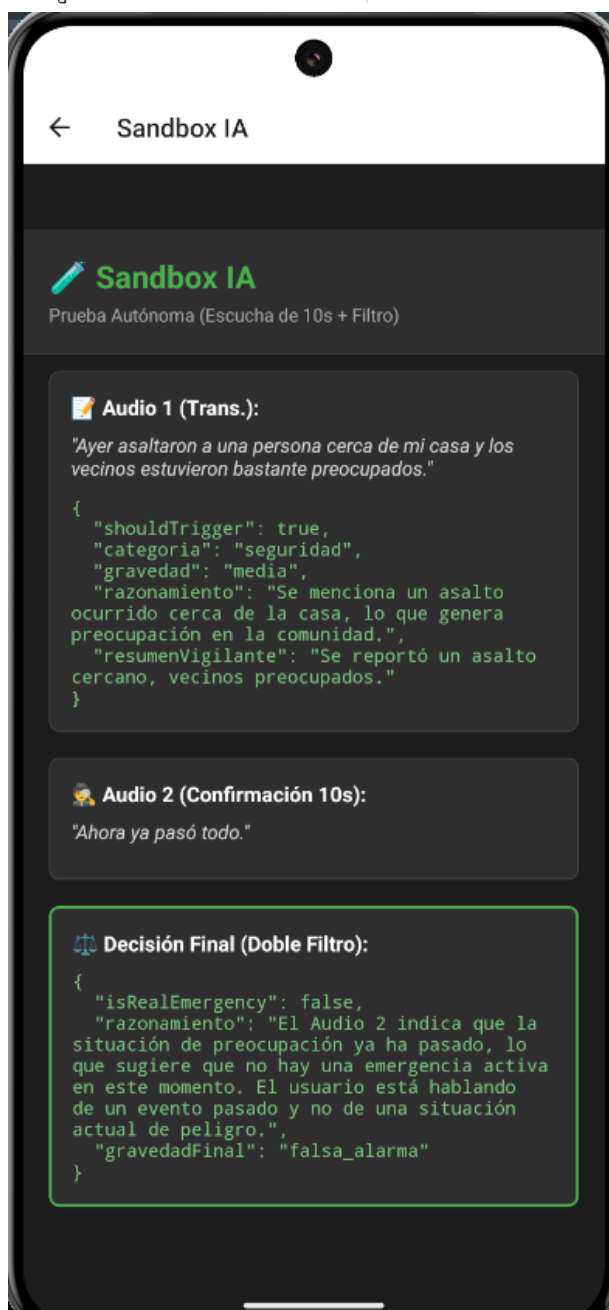
Caso 4:



Caso 5:



Caso 6:



Caso 7:

 Audio 1 (Trans.):

"Están hablando de un asalto que ocurrió esta mañana en el centro. Al parecer entraron a una tienda."

```
{
  "shouldTrigger": true,
  "categoria": "seguridad",
  "gravedad": "media",
  "razonamiento": "Se menciona un asalto
ocurrido, lo que indica una situación de
riesgo para la seguridad pública.",
  "resumenVigilante": "Se reporta un asalto
en el centro esta mañana."
}
```

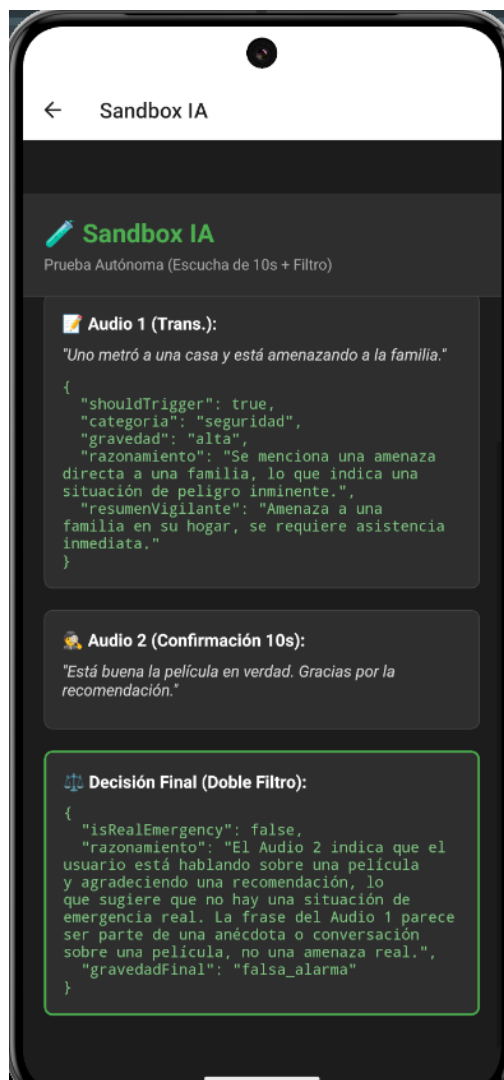
 Audio 2 (Confirmación 10s):

"Pero es solo una noticia si estoy viendo la tele. Acá no me ha pasado nada."

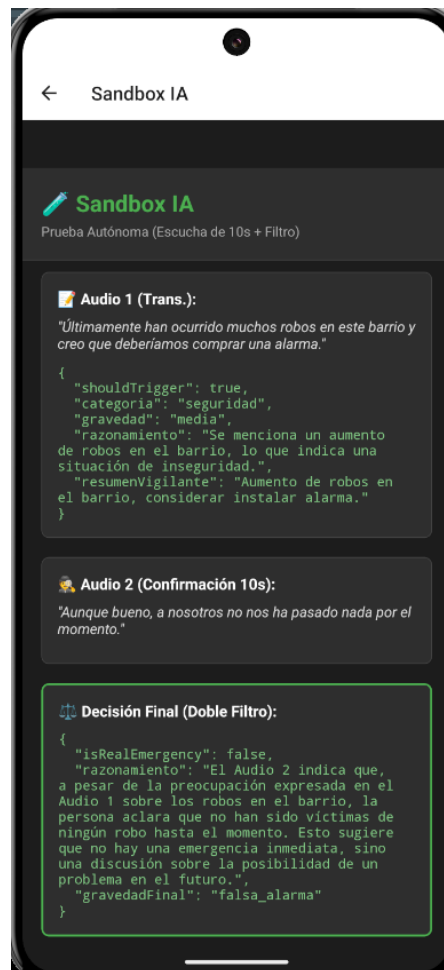
 Decisión Final (Doble Filtro):

```
{
  "isRealEmergency": false,
  "razonamiento": "El Audio 2 aclara que
el hablante está viendo una noticia sobre
un asalto en la televisión y que no está
involucrado en ninguna emergencia personal.
Esto indica que no hay una situación de
emergencia real.",
  "gravedadFinal": "falsa_alarma"
}
```

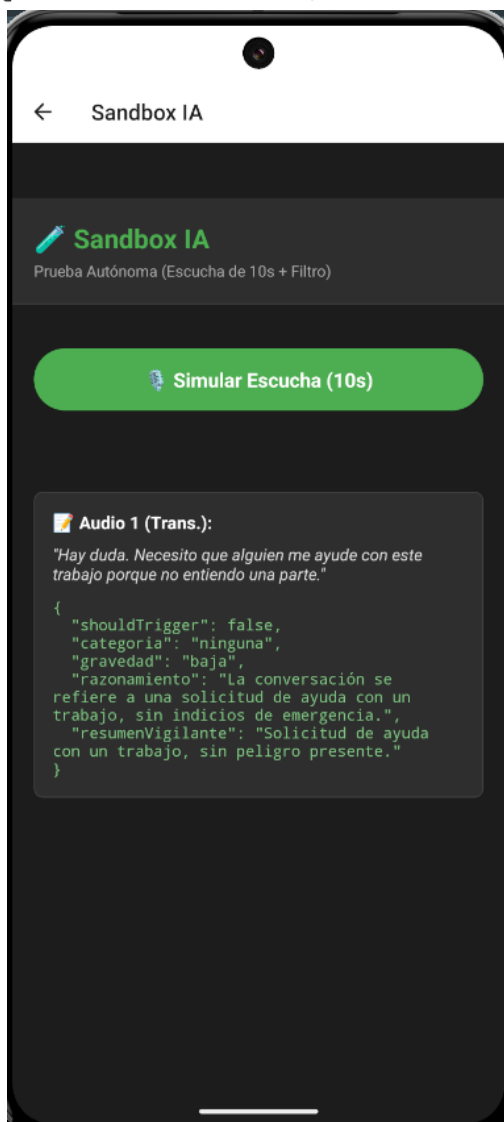
Caso 8:



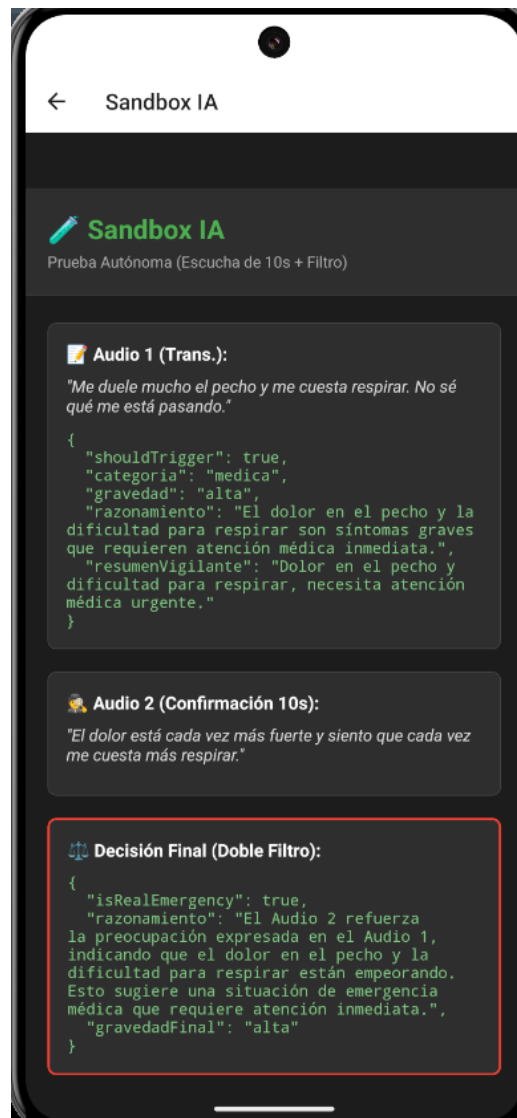
Caso 9:



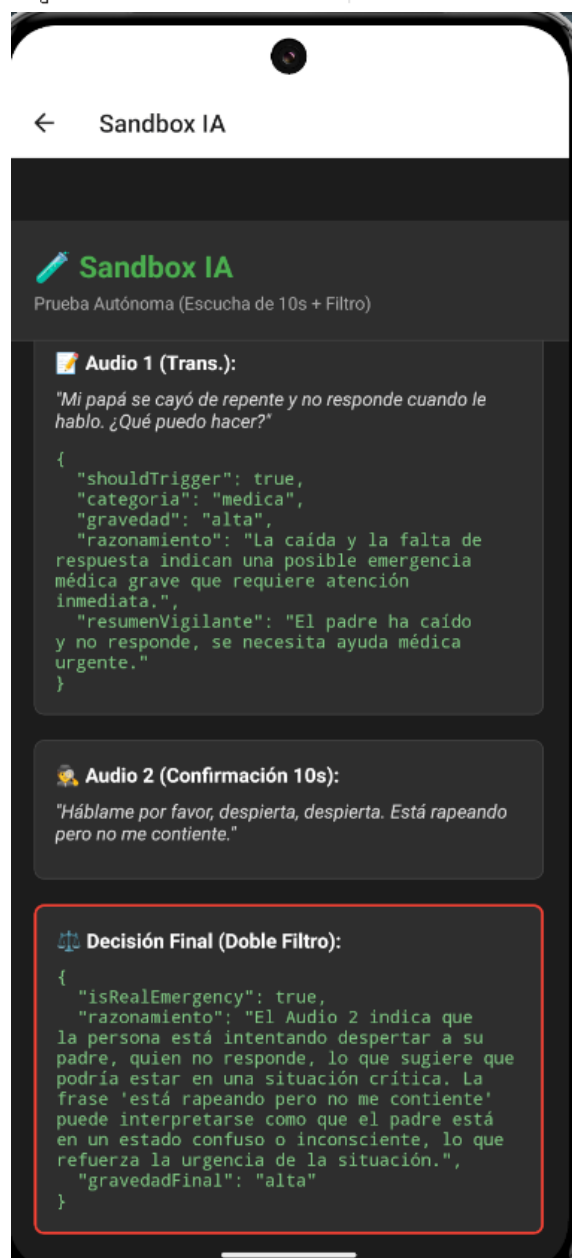
Caso 10:



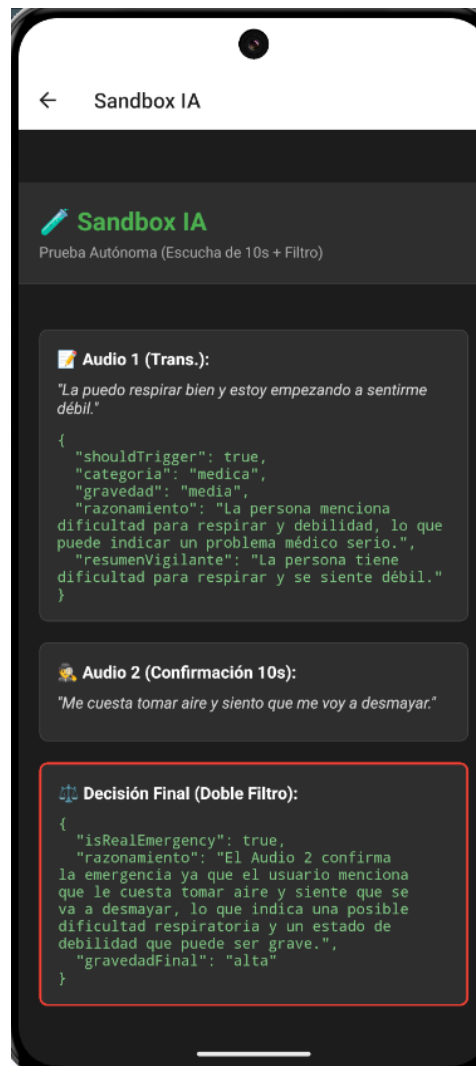
Caso 11:



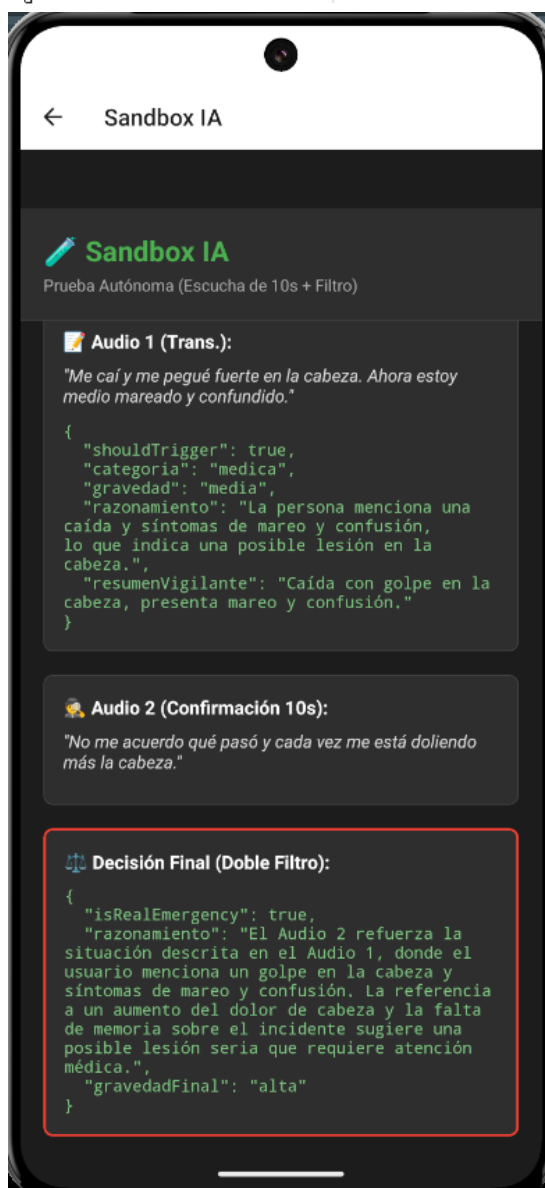
Caso 12:



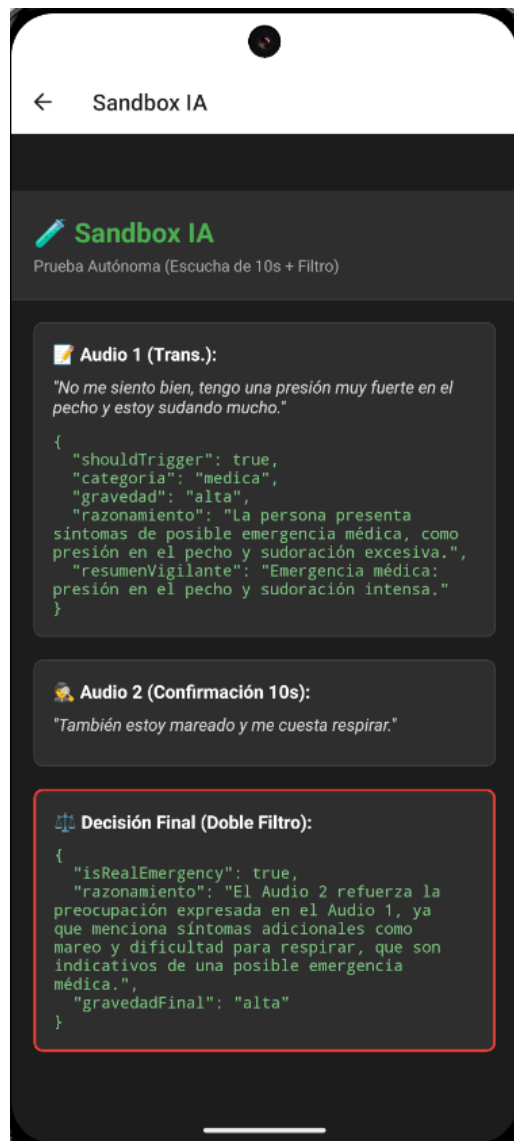
Caso 13:



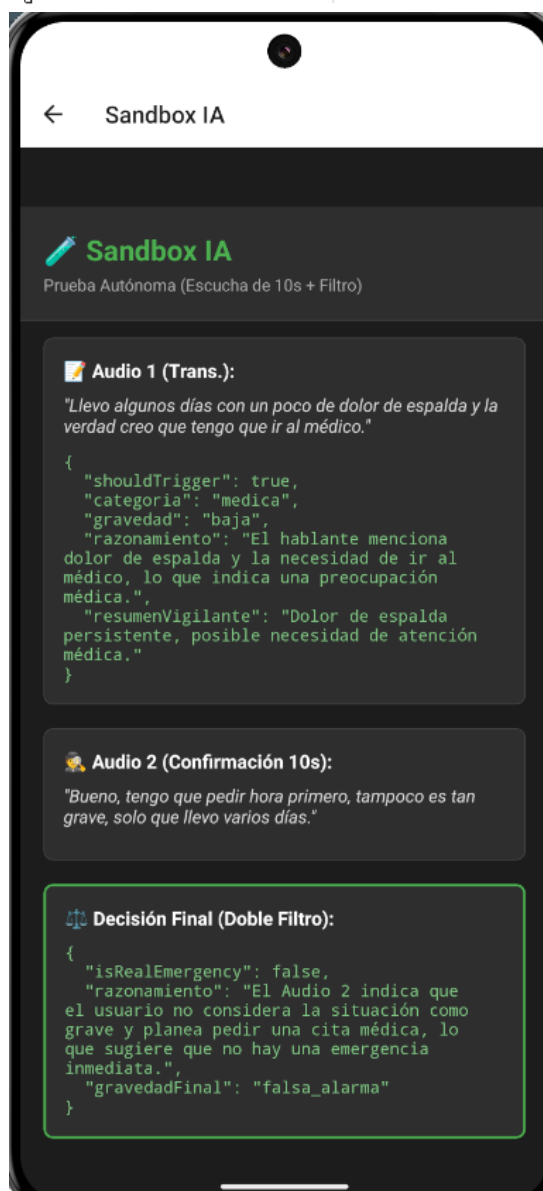
Caso 14:



Caso 15:



Caso 16:



Caso 17:



Caso 18:



← Sandbox IA

 Sandbox IA

Prueba Autónoma (Escucha de 10s + Filtro)

 Audio 1 (Trans.):

"Cuando era niño me enfermaba bastante sídio y una vez incluso tuve problemas para respirar."

```
{
  "shouldTrigger": true,
  "categoria": "medica",
  "gravedad": "media",
  "razonamiento": "Se menciona una enfermedad infantil y problemas para respirar, lo que indica una situación médica potencialmente seria.",
  "resumenVigilante": "Problemas respiratorios en la infancia, situación médica a considerar."
}
```

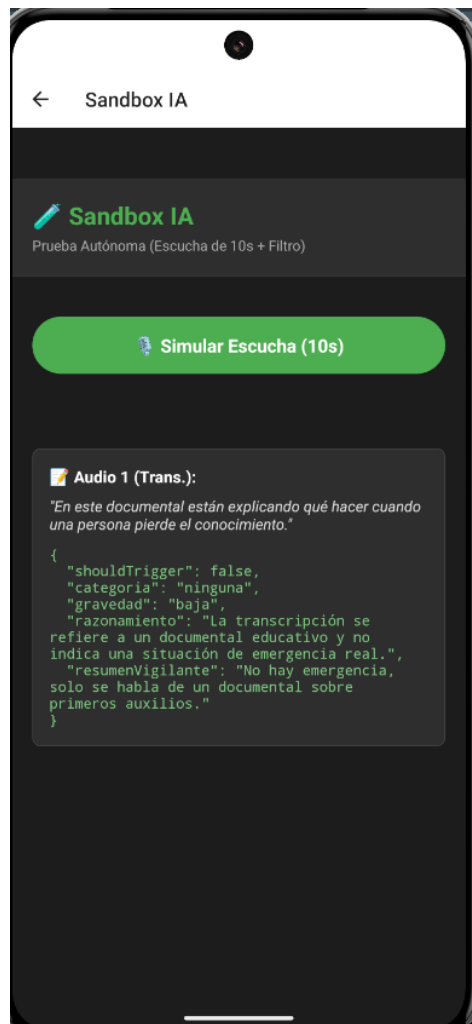
 Audio 2 (Confirmación 10s):

"Pero si ya fue hace rato, ahora estoy mejor y... y bueno."

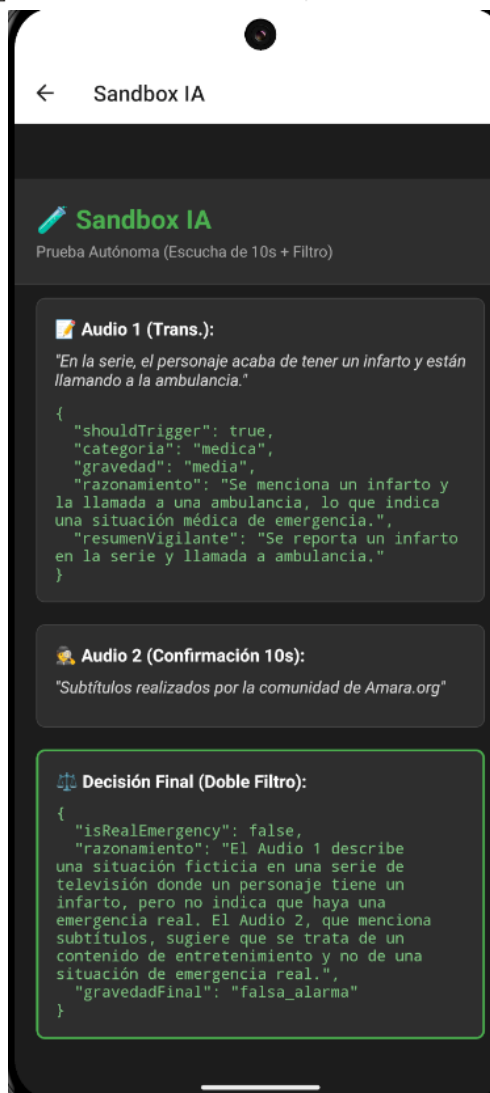
 Decisión Final (Doble Filtro):

```
{
  "isRealEmergency": false,
  "razonamiento": "El Audio 2 indica que la situación de problemas para respirar fue en el pasado y que el hablante está mejor ahora, lo que sugiere que no hay una emergencia actual.",
  "gravedadFinal": "falsa_alarma"
}
```

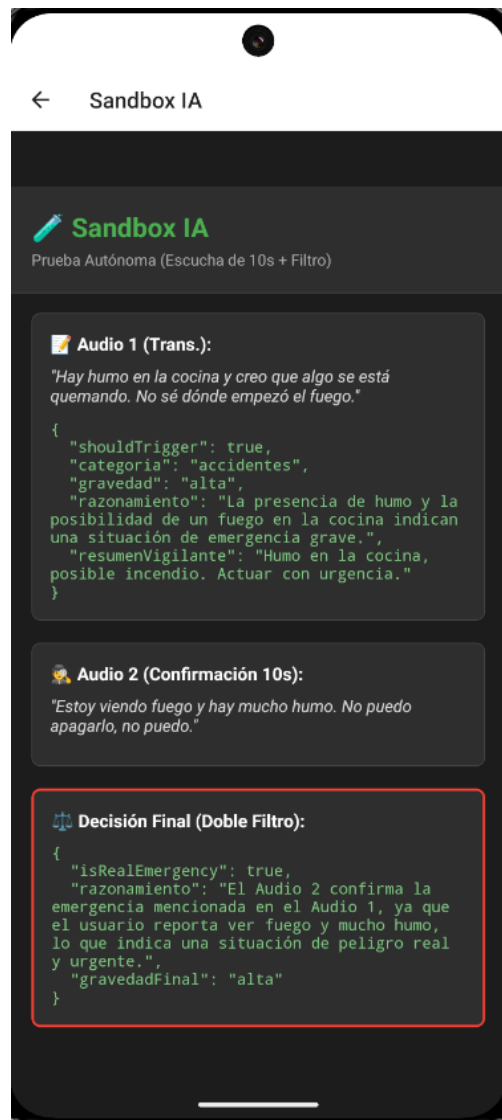
Caso 19:



Caso 20:



Caso 21:



Caso 22:



Caso 23:

← Sandbox IA

 **Sandbox IA**

Prueba Autónoma (Escucha de 10s + Filtro)

 **Audio 1 (Trans.):**

"Mi abuela se cayó en la escalera y no se puede levantar."

```
{
  "shouldTrigger": true,
  "categoria": "accidentes",
  "gravedad": "media",
  "razonamiento": "La caída de la abuela en la escalera indica un accidente que puede requerir atención médica.",
  "resumenVigilante": "Accidente: abuela cayó en la escalera y no puede levantarse."
}
```

 **Audio 2 (Confirmación 10s):**

"Me dice que le duele mucho la pierna y no puede apoyarse y no la puedo levantar."

 **Decisión Final (Doble Filtro):**

```
{
  "isRealEmergency": true,
  "razonamiento": "El Audio 2 confirma la emergencia mencionada en el Audio 1, ya que se detalla que la abuela tiene un dolor significativo en la pierna y no puede levantarse, lo que indica una situación de posible lesión que requiere atención médica.",
  "gravedadFinal": "alta"
}
```

Caso 24:



 Sandbox IA

 **Sandbox IA**
Prueba Autónoma (Escucha de 10s + Filtro)

 **Audio 1 (Trans.):**
"Me cayó agua hirviendo encima y me quemé el brazo."

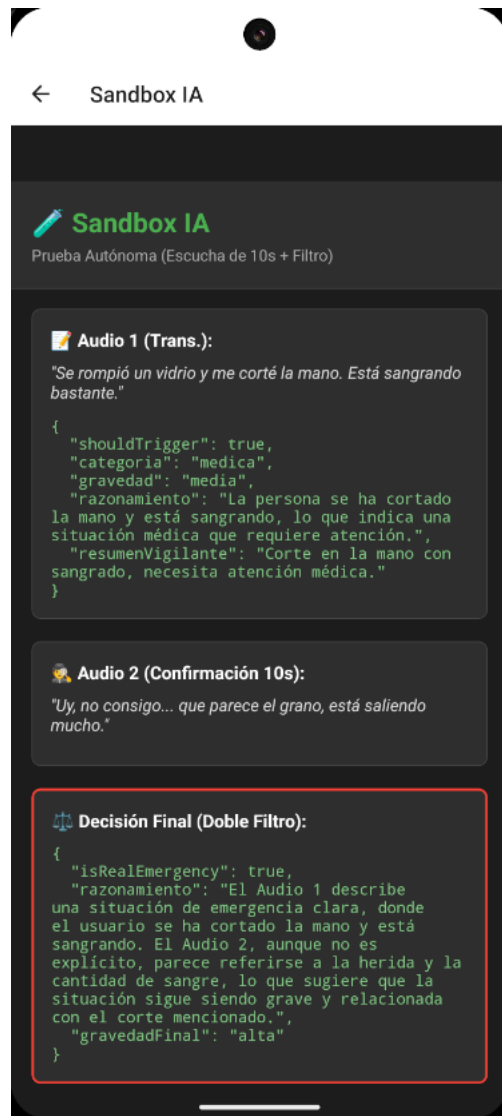
```
{
  "shouldTrigger": true,
  "categoria": "accidentes",
  "gravedad": "media",
  "razonamiento": "La persona ha sufrido una quemadura por agua hirviendo, lo que puede requerir atención médica.",
  "resumenVigilante": "Quemadura en el brazo por agua hirviendo, necesita atención médica."
}
```

 **Audio 2 (Confirmación 10s):**
"¡Ay! ¡Ay, ay, ay!"

 **Decisión Final (Doble Filtro):**

```
{
  "isRealEmergency": true,
  "razonamiento": "El Audio 1 describe una situación de quemadura por agua hirviendo, lo cual es una emergencia médica. El Audio 2, con los gritos de dolor, sugiere que la persona está sufriendo y confirma la gravedad de la situación.",
  "gravedadFinal": "alta"
}
```

Caso 25:



Caso 26:



← Sandbox IA

 Sandbox IA

Prueba Autónoma (Escucha de 10s + Filtro)

 Audio 1 (Trans.):

"Ayer me caí en la escalera y me hice un pequeño corte."

```
{
  "shouldTrigger": true,
  "categoria": "accidentes",
  "gravedad": "baja",
  "razonamiento": "La persona menciona una caída y un corte, lo que indica un accidente físico, aunque no parece grave.",
  "resumenVigilante": "Accidente menor: caída en la escalera con un pequeño corte."
}
```

 Audio 2 (Confirmación 10s):

"Lo bueno es que fue un rasguño nomás y bueno, ya pasó, ya fue."

 Decisión Final (Doble Filtro):

```
{
  "isRealEmergency": false,
  "razonamiento": "El Audio 2 indica que el incidente mencionado en el Audio 1 fue un rasguño menor y que ya ha pasado, lo que sugiere que no hay una emergencia actual.",
  "gravedadFinal": "falsa_alarma"
}
```

Caso 27:

 Sandbox IA

 **Sandbox IA**
Prueba Autónoma (Escucha de 10s + Filtro)

 **Audio 1 (Trans.):**
"Tengo una fuga debajo de la mano y necesito arreglarla."

```
{
  "shouldTrigger": true,
  "categoria": "accidentes",
  "gravedad": "media",
  "razonamiento": "La mención de una fuga sugiere un posible accidente que podría causar daño o peligro.",
  "resumenVigilante": "Fuga detectada, se necesita atención para evitar riesgos."
}
```

 **Audio 2 (Confirmación 10s):**
"Lo bueno es que solamente agua, así que nada, es que arreglarla nomás."

 **Decisión Final (Doble Filtro):**

```
{
  "isRealEmergency": false,
  "razonamiento": "El Audio 2 aclara que la fuga es solo de agua, lo que indica que no se trata de una emergencia grave. El usuario menciona que solo necesita arreglarla, lo que sugiere que es un problema menor y no una situación de emergencia.",
  "gravedadFinal": "falsa_alarma"
}
```

Caso 28:



Sandbox IA

Sandbox IA
Prueba Autónoma (Escucha de 10s + Filtro)

Audio 1 (Trans.):
"La casa al lado tuvo un incendio hace dos meses y todavía están reparando los daños."

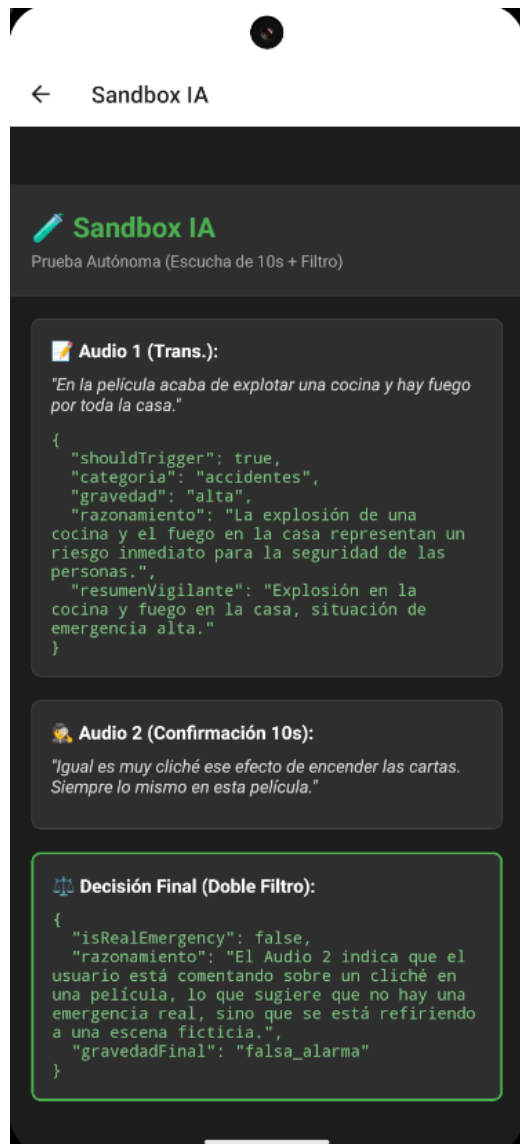
```
{
  "shouldTrigger": true,
  "categoria": "accidentes",
  "gravedad": "media",
  "razonamiento": "Se menciona un incendio en la casa vecina, lo que indica un riesgo potencial de accidentes similares.",
  "resumenVigilante": "Incendio en casa vecina hace dos meses, aún en reparaciones."
}
```

Audio 2 (Confirmación 10s):
"Bueno, ahora no hay fuego ni nada, solamente están haciendo las preparaciones, pero quedó a dar dolor a humo todavía."

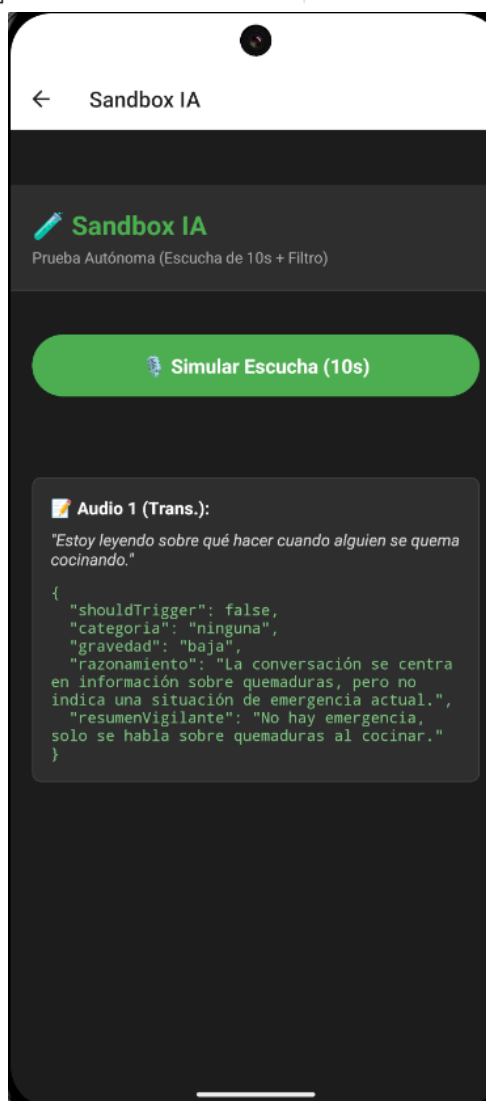
Decisión Final (Doble Filtro):

```
{
  "isRealEmergency": false,
  "razonamiento": "El Audio 2 indica que no hay fuego en la actualidad y que solo se están realizando reparaciones, lo que sugiere que no hay una emergencia activa en este momento. La mención de 'dolor a humo' se refiere a un efecto residual del incendio anterior, no a una situación de emergencia actual.",
  "gravedadFinal": "falsa_alarma"
}
```

Caso 29:

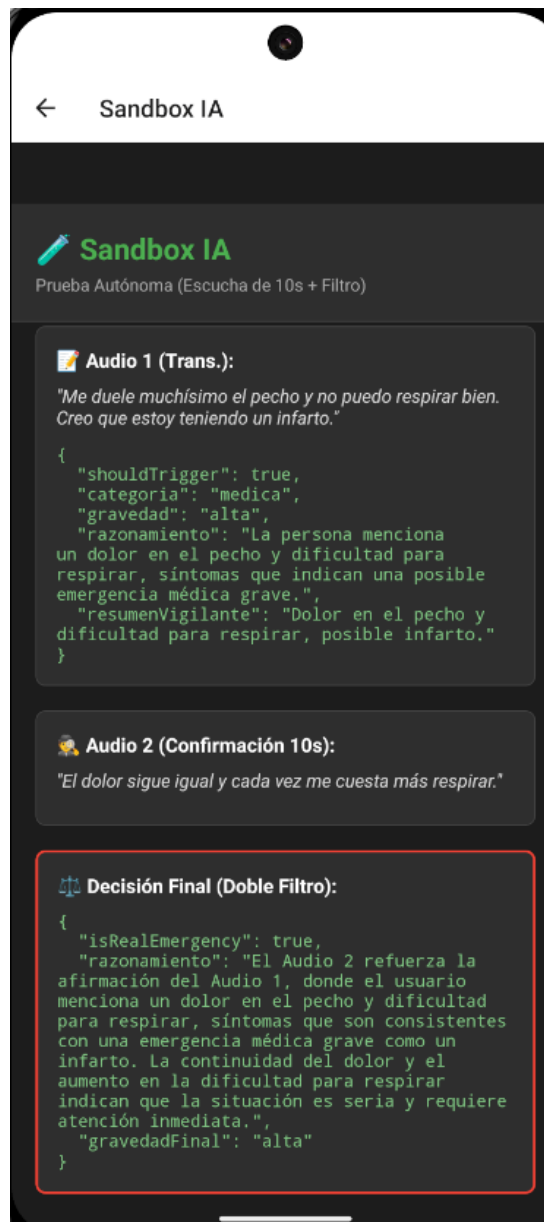


Caso 30:

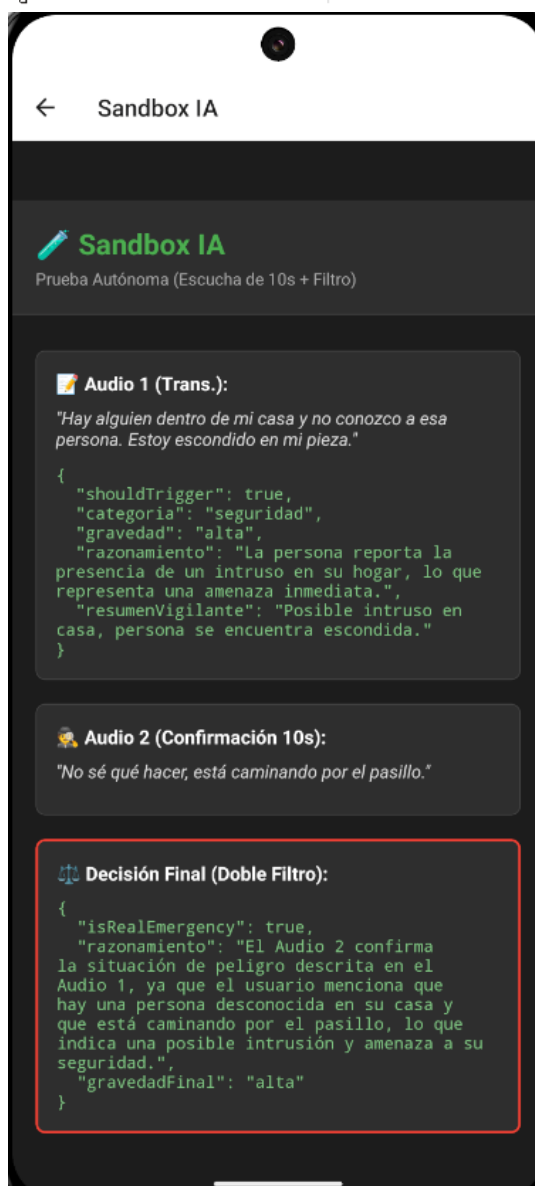


Anexo B. Evidencias de la segunda batería de validación

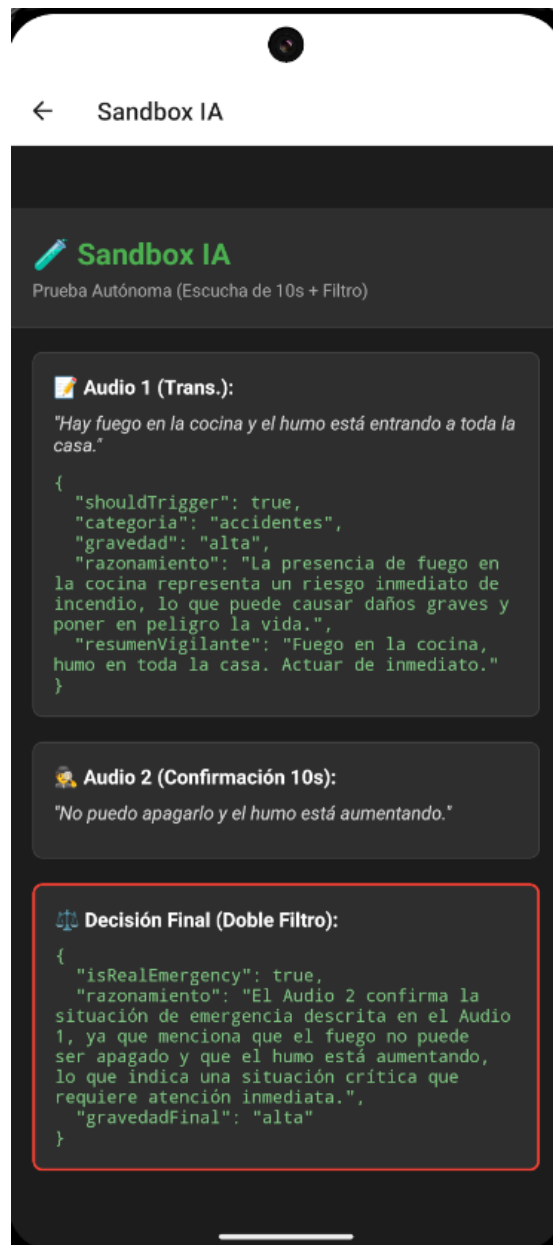
Caso 31:



Caso 32:



Caso 33:



Caso 34:



 Sandbox IA

 **Sandbox IA**
Prueba Autónoma (Escucha de 10s + Filtro)

 **Audio 1 (Trans.):**
"No puedo más. Necesito que alguien me ayude."

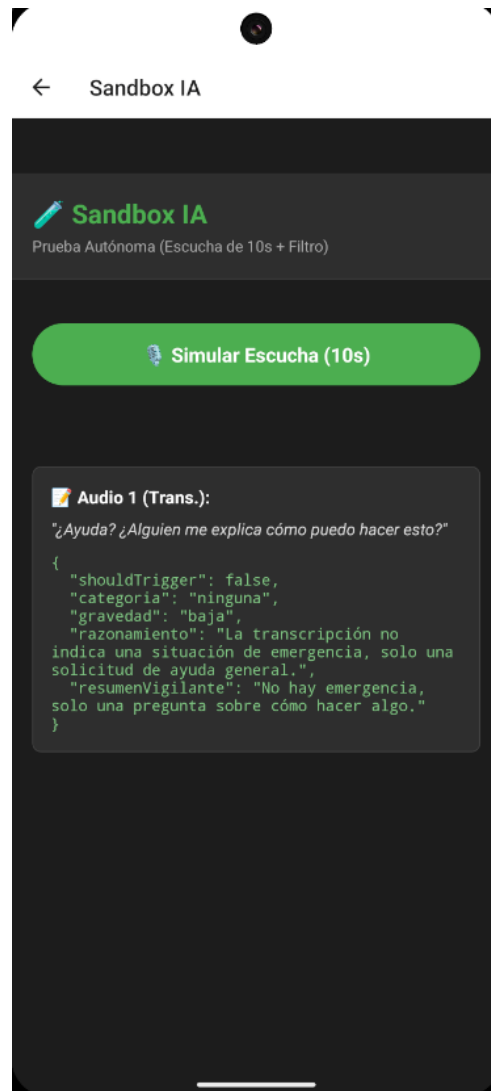
```
{
  "shouldTrigger": true,
  "categoria": "medica",
  "gravedad": "media",
  "razonamiento": "La frase indica una
necesidad urgente de ayuda, lo que sugiere
una posible situación médica.",
  "resumenVigilante": "Se necesita ayuda
urgente, posible situación médica."
}
```

 **Audio 2 (Confirmación 10s):**
*"Estoy tratando de terminar este trabajo y llevo toda la
noche haciéndolo."*

 **Decisión Final (Doble Filtro):**

```
{
  "isRealEmergency": false,
  "razonamiento": "El Audio 2 indica que
la persona está lidiando con una carga de
trabajo y no menciona ninguna situación de
peligro o emergencia. La frase del Audio
1 puede interpretarse como una expresión
de frustración relacionada con el trabajo,
lo que sugiere que no se trata de una
emergencia real.",
  "gravedadFinal": "falsa_alarma"
}
```

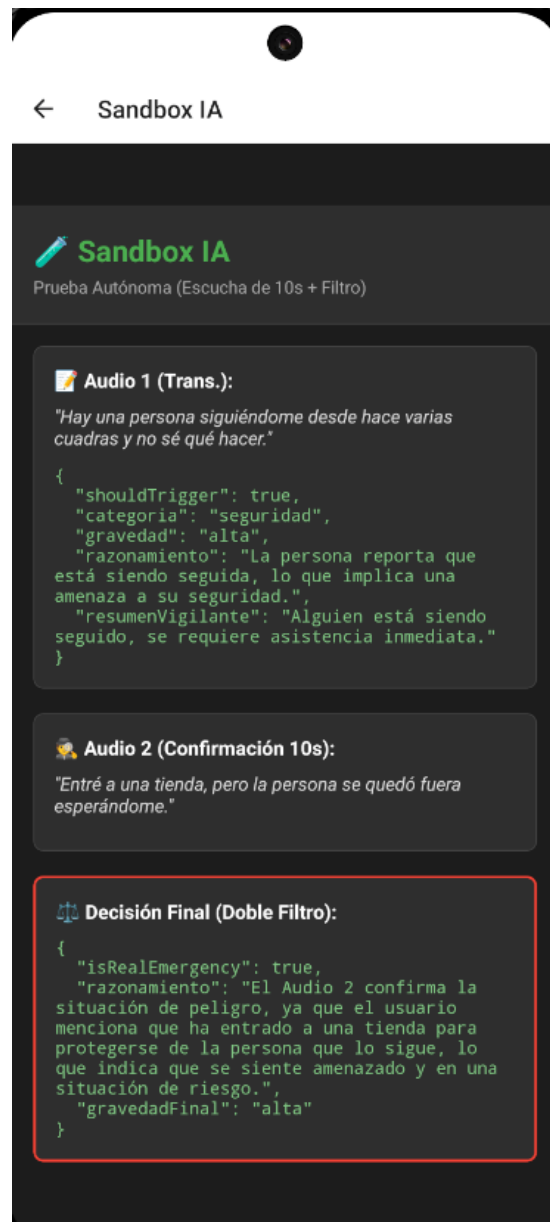
Caso 35:



Caso 36:



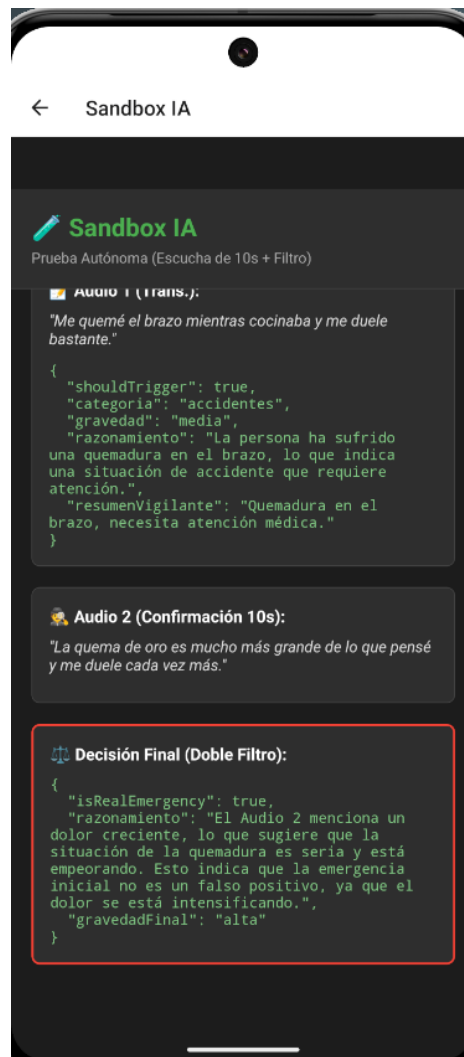
Caso 37:



Caso 38:



Caso 39:



Caso 40:



Anexo C. Planillas de resultados



Primera batería:

ID	Categoría	Tipo	Audio 1	Audio 2	Emergencia Esperada	Gravedad Esperada	Resultado Obtenido	Gravedad Obtenida	Razonamiento	¿Correcto?
1	Seguridad	Positivo	En casa. Escuché que abrió la puerta y ahora está caminando hacia mi casa. Escuché que abrió la puerta y ahora está caminando hacia mi casa.	Se está acercando a mi habitación y tengo miedo de que me encuentre.	TRUE	Alta	TRUE	Alta	Acción por la presencia de una persona en la habitación.	TRUE
2	Seguridad	Positivo	Estoy comiendo y está diciendo que me va a hacer daño. No me da miedo.	La puerta y sigue amenazándome. No puedo salir y necesito que alguien me ayude.	TRUE	Alta	TRUE	Alta	Se siente atrapado y necesita ayuda.	TRUE
3	Seguridad	Positivo	Hay una persona que viene caminando detrás de mí desde la puerta. Creo que está tratando de forzar la cerradura. Estoy escondido.	Hay una tienda para esconderme, pero está esperando afuera y no puedo salir.	TRUE	Media/Alta	TRUE	Media(audio 1)/Alta (audio 2)	Lo que se va a la cerradura.	TRUE
4	Seguridad	Positivo	mi casa y hay dos personas intentando abrir la puerta. Creo que están tratando de forzar la cerradura. Estoy escondido.	Hay una tienda para esconderme, pero está esperando afuera y no puedo salir.	TRUE	Alta	TRUE	Alta	Alta	TRUE
5	Seguridad	Positivo	No sé si debería llamar a alguien. Estoy encerrado en una pieza. Está golpeando la puerta y gritando que va a entrar. No quiero salir.	Estoy encerrado en una pieza. Está golpeando la puerta y gritando que va a entrar. No quiero salir.	TRUE	Alta	TRUE	Media(audio 1)/Alta (audio 2)	persona alterada está gritando.	TRUE
6	Ninguna	Negativo	Ayer asaltaron a una persona cerca de mi casa y los vecinos. Ahora ya pasó todo, la policía vino ayer y actualmente está todo tranquilo.	Es una noticia que estoy viendo en televisión, no está pasando nada.	FALSE	Falsa alarma	FALSE	Falsa alarma	emergencia activa en este momento.	TRUE
7	Ninguna	Negativo	Están hablando de un asalto que ocurrió esta mañana en el centro. Es una noticia que estoy viendo en televisión, no está pasando nada.	Es una noticia que estoy viendo en televisión, no está pasando nada.	FALSE	Falsa alarma	FALSE	Falsa alarma	no está involucrado.	TRUE
8	Ninguna	Negativo	película un hombre entró a una casa y está amenazando a la familia. Es una película, estoy sentado viendo televisión. No está ocurriendo nada.	Es una película, estoy sentado viendo televisión. No está ocurriendo nada.	FALSE	Falsa alarma	FALSE	Falsa alarma	es una situación de emergencia.	TRUE
9	Ninguna	Negativo	jurado muchos robos en este barrio y creo que deberíamos estar más alerta. No está pasando nada, solamente estamos conversando sobre el tema.	Es una noticia que estoy viendo en televisión, no está pasando nada.	FALSE	Falsa alarma	FALSE	Falsa alarma	temas de ningún robo.	TRUE
10	Ninguna	Negativo	sito que alguien me ayude con este trabajo porque no entiendo nada. Es el informe de la universidad que tengo que entregar mañana.	Es el informe de la universidad que tengo que entregar mañana.	FALSE	Falsa alarma	FALSE	Falsa alarma	ciudad de ayuda con un problema.	TRUE
11	Médica	Positivo	Me duele muchísimo el pecho y me cuesta respirar. No sé qué está pasando. El dolor está empeorando y siento que me voy a desmayar. Cada vez me cuesta más respirar.	Me duele muchísimo el pecho y me cuesta respirar. No sé qué está pasando. El dolor está empeorando y siento que me voy a desmayar. Cada vez me cuesta más respirar.	TRUE	Alta	TRUE	Alta	habilidad para respirar.	TRUE
12	Médica	Positivo	Mi papá se cayó de repente y no responde cuando le hablo. Sigue sin responder y no puedo despertarlo. Está respirando, pero no responde.	Sigue sin responder y no puedo despertarlo. Está respirando, pero no responde.	TRUE	Alta	TRUE	Alta	La frase "está respirando" indica que se va a desmayar.	TRUE
13	Médica	Positivo	No puedo respirar bien y estoy empezando a sentirme muy débil. Me cuesta mucho tomar aire y siento que me voy a desmayar. Cada vez me cuesta más respirar.	Me cuesta mucho tomar aire y siento que me voy a desmayar. Cada vez me cuesta más respirar.	TRUE	Alta	TRUE	Alta	que se va a desmayar.	TRUE
14	Médica	Positivo	Me caí y me golpeé muy fuerte la cabeza. Ahora estoy mareado. No recuerdo bien qué pasó y el dolor de cabeza está aumentando. Estoy confundido.	No recuerdo bien qué pasó y el dolor de cabeza está aumentando. Estoy confundido.	TRUE	Alta	TRUE	Media(audio 1)/Alta (audio 2)	ción. La referencia a un accidente.	TRUE
15	Médica	Positivo	No me siento bien, tengo una presión muy fuerte en el pecho. Ahora también estoy mareado y me cuesta respirar. Creo que necesito ayuda.	Ahora también estoy mareado y me cuesta respirar. Creo que necesito ayuda.	TRUE	Alta	TRUE	Alta	significativos como mi estado.	TRUE
16	Ninguna	Negativo	Llevo algunos días con un poco de dolor de espalda y creo que no es algo urgente, puedo seguir haciendo mis cosas normalmente.	No es algo urgente, puedo seguir haciendo mis cosas normalmente.	FALSE	Falsa alarma	FALSE	Falsa alarma	me y planes pedir una cita.	TRUE
17	Ninguna	Negativo	La semana pasada mi abuelo se desmayó y tuvimos que llevarlo al hospital. Pero ya está en la casa, se recuperó y ahora se encuentra bien.	Pero ya está en la casa, se recuperó y ahora se encuentra bien.	FALSE	Falsa alarma	FALSE	Falsa alarma	caso, lo que sugiere que el paciente está mejor.	TRUE
18	Ninguna	Negativo	Cuando era niño me enfermaba bastante seguido y una vez tuve que ir al hospital. Ahora estoy bien y solamente estoy contando los días que no estoy enfermo.	hace muchos años. Ahora estoy bien y solamente estoy contando los días que no estoy enfermo.	FALSE	Falsa alarma	FALSE	Falsa alarma	pasado y que el habla.	TRUE
19	Ninguna	Negativo	En la serie el personaje acaba de tener un infarto y están llamando a un médico. Estoy viendo la serie en televisión. Yo estoy bien y no necesito ayuda.	Estoy viendo la serie en televisión. Yo estoy bien y no necesito ayuda.	FALSE	Falsa alarma	FALSE	Falsa alarma	don donde un personaje.	TRUE
20	Ninguna	Negativo	En este documental están explicando qué hacer cuando una persona tiene un accidente. Es una información muy útil que estoy viendo.	Es una información muy útil que estoy viendo.	FALSE	Falsa alarma	FALSE	Falsa alarma	intal educativo y no informativo.	TRUE
21	Accidente	Positivo	Hay humo en la cocina y creo que algo se está quemando. No quiero ir a ver, pero necesito ayuda. El olor es fuerte.	Hay humo en la cocina y creo que algo se está quemando. No quiero ir a ver, pero necesito ayuda. El olor es fuerte.	TRUE	Alta	TRUE	Alta	usuario reporta ver fuego.	TRUE
22	Accidente	Positivo	Hay un olor muy fuerte a gas en la cocina y no sé de dónde viene. El olor está cada vez más fuerte y creo que hay una fuga. Estoy muy preocupado.	El olor está cada vez más fuerte y creo que hay una fuga. Estoy muy preocupado.	TRUE	Alta	TRUE	Media(audio 1)/Alta (audio 2)	cha de una fuga, lo que indica un peligro.	TRUE
23	Accidente	Positivo	Mi abuela se cayó en las escaleras y no puede levantarse. Le duele mucho la pierna y no puede apoyarse. Necesitamos ayuda.	Le duele mucho la pierna y no puede apoyarse. Necesitamos ayuda.	TRUE	Media/Alta	TRUE	Media(audio 1)/Alta (audio 2)	por significativo en la historia.	TRUE
24	Accidente/Médica	Positivo	Se me cayó agua hirviendo encima y me quemé el brazo. Me duele muchísimo y la quemadura ocupa una parte grande de mi brazo.	Me duele muchísimo y la quemadura ocupa una parte grande de mi brazo.	TRUE	Media/Alta	TRUE	Media(audio 1)/Alta (audio 2)	édica. El audio 2, continúa la descripción.	TRUE
25	Accidente/Médica	Positivo	Se rompió un vidrio y me corté la mano. Está sangrando bastante. La sangre no deja de salir y no consigo detener el sangrado.	La sangre no deja de salir y no consigo detener el sangrado.	TRUE	Alta	TRUE	Media(audio 1)/Alta (audio 2)	o es explícito, parece que hay una emergencia.	TRUE
26	Ninguna	Negativo	Ayer me caí en las escaleras y me hice un pequeño corte. Pero ya estoy bien, fue solamente un rasguño y ahora estoy haciendo mis cosas normalmente.	Pero ya estoy bien, fue solamente un rasguño y ahora estoy haciendo mis cosas normalmente.	FALSE	Falsa alarma	FALSE	Falsa alarma	un rasguño menor y no una emergencia.	TRUE
27	Ninguna	Negativo	Tengo una fuga debajo del lavamanos y necesito arreglarla. Es solamente agua, lleva varios días goteando y no hay ningún problema.	Es solamente agua, lleva varios días goteando y no hay ningún problema.	FALSE	Falsa alarma	FALSE	Falsa alarma	usuario menciona que no es una emergencia.	TRUE
28	Ninguna	Negativo	La casa de al lado tuvo un incendio hace dos meses y todavía ahora no hay fuego ni humo. Solamente están haciendo los repuestos.	Ahora no hay fuego ni humo. Solamente están haciendo los repuestos.	FALSE	Falsa alarma	FALSE	Falsa alarma	encia activa en este momento.	TRUE
29	Ninguna	Negativo	En la película acaba de explotar una cocina y hay fuego por todos lados. Es solamente la película que estoy viendo. En mi casa está todo tranquilo.	Es solamente la película que estoy viendo. En mi casa está todo tranquilo.	FALSE	Falsa alarma	FALSE	Falsa alarma	ula, lo que sugiere que no hay una emergencia.	TRUE
30	Ninguna	Negativo	Estoy leyendo sobre qué hacer cuando alguien se quema con agua caliente. Es para una tarea de primeros auxilios. No hay nadie herido, solo estoy leyendo.	Es para una tarea de primeros auxilios. No hay nadie herido, solo estoy leyendo.	FALSE	Falsa alarma	FALSE	Falsa alarma	re quemaduras, pero no es una emergencia.	TRUE

Segunda batería:

CASO	Categoría	Tipo	Audio 1	Resultado Audio 1	Gravedad Audio 1	¿Requería Audio 2?	Audio 2	Resultado final	Gravedad Final	¿Cambió la gravedad?	¿Decisión final correcta?	Razonamiento	Latencia Audio 1	Latencia Audio 2	Latencia total
31	Médica	Positivo	Transcripción	Emergencia	Alta	Si	Transcripción	Emergencia	Alta	No	Si	emergencia médica	4.11	5.67	9.78
32	Seguridad	Positivo	Transcripción	"Sospecha"	Alta	Si	Transcripción	Emergencia	Alta	No	Si	no desconocido	4.88	5.85	10.73
33	Accidente	Positivo	Transcripción	Emergencia	Alta	Si	Transcripción	Emergencia	Alta	No	Si	puede ser apagada	4.63	3.75	8.38
34	Prueba al sistema	Negativo	Transcripción	Emergencia	Media	Si	Transcripción	Falsa Alarma	No hay	Si	Si	se del Audio 1	3.6	2.26	5.86
35	Prueba al sistema	Negativo	Transcripción	Ninguna	Baja	No	Nada	Falsa Alarma	No hay	No hubo segundo	Si	en de emergencia	3.92	0	3.92
36	Prueba al sistema	Negativo	Transcripción	Sospecha	Alta	Si	Transcripción	Falsa Alarma	No hay	Si	Si	idente mencionado	4.22	2.15	6.37
37	Seguridad	Positivo	Transcripción	Sospecha	Alta	Si	Transcripción	Emergencia	Alta	No	Si	la para proteger	5.14	5.33	10.47
38	Médica	Positivo	Transcripción	Sospecha	Media	Si	Transcripción	Emergencia	Alta	Si	Si	o puede mantener	3.71	2.49	6.2
39	Accidente	Positivo	Transcripción	Emergencia	Media	Si	Transcripción	Emergencia	Alta	Si	Si	está empeorando	3.31	3.96	7.27
40	Prueba al sistema	Negativo	Transcripción	Emergencia	Alta	Si	Transcripción	Falsa Alarma	No hay	Si	Si	elícula y no a un	3.5	2.31	5.81
													4.102	3.377	7.479