



TRABAJO FINAL DE MÁSTER

IA EN LA GESTIÓN DE CITAS MÉDICAS: INNOVANDO EN
CONFIRMACIÓN Y REDUCCIÓN DE AUSENCIAS.

Descripción breve

La utilización de inteligencia artificial (IA) en la gestión de citas médicas tiene como objetivo mejorar la asignación de estas, disminuyendo los “non show” con el objetivo de mejorar la eficiencia de los servicios de salud ambulatorios

Grupo 3

Hernando Acevedo Aguilar
Michelle Alexandra Chicaiza Anrrango
Luis Marcelo Ortiz Carinao
Sergio Valdueza Lozano

Resumen ejecutivo del proyecto

Uno de los principales problemas de los servicios de salud tanto de España como de Latinoamérica son los tiempos de espera para consultas médicas de especialidades. En España para acceder a una atención especializada, los pacientes esperan una media de 79 días para ser atendidos en primera consulta con un rango entre 22 y 107 días (1). Mientras que el sistema de salud chileno enfrenta un desafío crítico que pese a tener patologías priorizadas por ley, donde se establece un tiempo de espera máximo, (garantías explícitas de salud GES) se alcanzaron 61.191 garantías retrasadas al 31 de diciembre del 2022, donde cada garantía promediaba un retraso de 156.5 días (2).

Esto se explica por múltiples factores, tales como la brecha de especialistas (3), y la **inasistencia a consultas médicas, que en algunos centros alcanza hasta el 20%** (4) y (5).

Debido a esta problemática es que planteamos el desafío de integrar herramientas de inteligencia artificial con el objetivo de mejorar la asignación de citas médicas, mediante la predicción de una probabilidad de inasistencia (“non show”) con herramientas de machine learning, y en base a esta probabilidad de inasistencia ofrecer estas horas a pacientes en listas de espera con el fin de disminuir los retrasos en las consultas médicas de especialidades.

Para este proyecto se plantean una serie de objetivos e hitos clave en su desarrollo:

- 1) Se obtiene una base de datos con información de cada usuario del servicio de salud y una etiqueta target: asistió o no a su cita.
- 2) Se realiza un análisis de los datos, donde se identifican y preparan variables predictoras de inasistencia.
- 3) Se preparan los datos para su posterior uso con herramientas de machine learning mediante estandarización, balanceo de clases y técnicas de reducción de dimensionalidad.
- 4) Se entrenan modelos de machine learning, en donde previa separación de un conjunto de datos de prueba se valorará la capacidad predictora de inasistencia (“non show”) de estos modelos
- 5) En base a la información obtenida del punto anterior se elaborará un modelo de asignación de cupos de “overbooking” con el fin de minimizar el problema de que un “non show” signifique una pérdida de horas médicas, pero a su vez que evite de que haya más de un paciente asignado para la misma hora médica, evitando así los retrasos en las consultas médicas. Vale decir corresponde a un problema de optimizar una función.
- 6) * Se integra este análisis de probabilidad de inasistencia y “overbooking” con una base de datos integrada a una interfaz de usuario.

Para la obtención de datos se recopilan fuentes internas de la empresa y se procede a realizar un análisis de bases de datos de fuentes externas que se encuentren con la variable objetivo: “asiste”, “non show” etiquetada.

Índice

- Resumen ejecutivo
- Índice
- Introducción
- Bibliografía

BORRADOR

Introducción

- breve introducción de la empresa: sector al que pertenece alcance geográfico y actividad desarrollada.

MediAgenda Solutions, S.L. (inventada para fines académicos), es una empresa líder en el sector de la salud, especializada en la administración inteligente de agendas médicas para hospitales y centros médicos en España y Latinoamérica. Nuestra actividad se centra en gestionar las agendas de visita a los especialistas, con la misión de optimizar la asignación de horarios médicos, y así reducir los tiempos de espera y mejorar la calidad de la atención médica.

- definición del problema planteado

En España, para acceder a una atención especializada, los pacientes esperan una media de 79 días hasta ser atendidos en primera consulta, con un rango entre 22 y 107 días (1).

Y en Chile, por ejemplo, el sistema de salud enfrenta un desafío crítico ya que, pese a tener patologías priorizadas por ley donde se establece un tiempo de espera máximo (garantías explícitas de salud GES), al 31 de diciembre del 2022, se alcanzaron 61.191 garantías retrasadas, donde cada garantía promediaba un retraso de 156.5 días (2). Esto se explica por múltiples factores, tales como la brecha de especialistas (3) y la **inasistencia a consultas médicas, que en algunos centros alcanza hasta el 20%** (4) y (5).

Éste es un obstáculo recurrente que entorpece nuestro objetivo: **la falta de asistencia de los pacientes a sus citas programadas**. Este alto porcentaje de “No Show” genera un desequilibrio en la programación de los médicos, afectando negativamente la eficiencia operativa y generando largas listas de espera para ser atendido por un especialista.

- objetivo del proyecto según el reto planteado

Debido a esta problemática, planteamos el desafío de integrar herramientas de inteligencia artificial para mejorar la asignación de citas médicas, mediante la predicción de una probabilidad estadística de inasistencia (“non show”) con herramientas de machine learning y, en base a esta probabilidad, optimizar las agendas médicas con un sistema de overbooking, similar al realizado en otros sectores (aviación, hoteles, etc.), con el fin de ofrecer estas horas a pacientes en listas de espera y disminuir así los retrasos en las consultas médicas de especialistas.

Adicional, queremos crear una aplicación de contacto entre centros médicos y pacientes que permita la concertación y confirmación de citas de forma eficiente, intuitiva y amigable. Queremos que la experiencia de agendar una consulta sea fluida y beneficiosa para todos los involucrados, aprovechando la tecnología para asignar citas, generar recordatorios, y, a su vez, recopilar más información que pueda ser útil a posteriori para realizar mejores predicciones de asistencia en nuestro modelo de IA.

- presentación resumida de las fases en las que va a contar el proyecto

Para la ejecución de este proyecto se plantean una serie de fases y sub-fases que nos ayuden a conseguir el objetivo planteado:

- **Fase 1:** Realizar un modelo de IA capaz de predecir la Asistencia o No Asistencia de los pacientes a las citas médicas:
 - Obtener una base de datos con información de las citas programadas a un servicio de salud, así como de los usuarios que las concertaron, y una etiqueta objetivo: asistió o no a su cita.
 - Realizar un Análisis Exploratorio de los Datos (EDA, del inglés “Exploratory Data Analysis”, donde se identifican y preparan las variables predictoras de inasistencia del dataset original, haciendo una limpieza de datos para tratar inconsistencias o valores faltantes, así como explorando la base de datos con una mente abierta en busca de relaciones, tendencias o anomalías.
 - Realizar Feature Engineering para obtener un dataset definitivo:
 - Extracción de características, transformando los datos brutos de la base de datos en características significativas.
 - Creación de nuevas características, combinando las existentes o trayéndolas de otras bases de datos (información climática, por ejemplo).
 - Selección de características, identificando las relevantes y eliminando las menos correlacionadas con la variable objetivo.
 - Manejo de datos faltantes.
 - Estandarización o Normalización de las características.
 - Reducción de la Dimensionalidad, aplicando técnicas como PCA o t-SNE.
 - Manejo de Variables Categóricas, creando columnas binarias para cada categoría (One-Hot Encoding) o asignando enteros únicos a cada categoría (Label Encoding).
 - Balanceo de Clases, usando técnicas de Oversampling, Undersampling, o SMOTE, que combina ambas técnicas anteriores.
 - Entrenar diferentes modelos de Machine Learning. Para ello primero se separa el dataset en datos de entrenamiento y datos de prueba. Y, a su vez, del conjunto de datos de entrenamiento se separa otro menor que sirva de validación y podamos usar para poder seleccionar los mejores hiperparámetros de cada modelo valorado. Se comparará y valorará la capacidad predictora de inasistencia (“No Show”) de cada modelo.
- **Fase 2:** Desarrollar uno o varios modelos de asignación de citas médicas que, utilizando las predicciones de asistencia conseguidas en la fase anterior, y, mediante la creación de cupos de overbooking, maximice el uso de los recursos médicos disponibles, tanto humanos (doctores especialistas) como materiales (instalaciones). Lo anterior sin generar tiempos de espera excesivos para los pacientes, ni una sobrecarga adicional en los servicios médicos fuera de horario normal. Estos dos últimos efectos son daños colaterales de overbooking, y por regla general son inversamente proporcionales a un uso más productivo de los recursos.

Así pues, los modelos son problemas de optimización donde minimizaremos una Función de Coste, la cual mide el costo de tener slots sin visita médica, el de

tener al doctor trabajando horas extra o el del paciente esperando a ser atendido, entre otros. Dicha función vendrá definida por unos factores ambientales (predicciones de asistencia, número de slots por día, etc.) y unas reglas de cita o restricciones (número máximo de pacientes programados para un único slot, sobrecarga máxima de atención al paciente, etc.). El objetivo del modelo a desarrollar es que, dados los parámetros que identifiquen dichos factores y reglas anteriores, aprenda el grado de overbooking a ocupar para, asignando las citas médicas en estricto orden de solicitud, se minimice el coste calculado.

- **Fase 3:** Creación de un Asistente Virtual basado en el Procesamiento de Lenguaje Natural (NLP) que permita el contacto entre Centros de Salud y Pacientes para la gestión de citas y recordatorios.

En esta fase, se desarrollará un Asistente Virtual que permitirá la comunicación entre los Centros de Salud y pacientes para la gestión de citas y recordatorios. El objetivo principal es mejorar la experiencia del usuario y optimizar el proceso de programación y seguimiento de citas médicas.

El Asistente Virtual se basará en técnicas de Procesamiento de Lenguaje Natural (NLP) para comprender las consultas y solicitudes de los pacientes, y proporcionar respuestas relevantes y precisas.

Fase 4: Desarrollar, *a nivel teórico*, un proceso de refinamiento del sistema que integre la base de datos de la interfaz creada para gestionar las citas y los recordatorios a los pacientes (Fase 3) con los modelos predictivo y de optimización generados en las Fases 1 y 2.

- Selección de posibles nuevas características relevantes provenientes de la aplicación.
- Adhesión de estas nuevas características predictivas al dataset definitivo elaborado en la Fase 1.
- Desarrollo fine-tuning del modelo de IA seleccionado en la Fase 1, con el objetivo de mejorar en precisión las predicciones de asistencia.
- Integrar los nuevos resultados en modelo de IA creado en la Fase 2, analizando la optimización / disminución de “coste” obtenida.
- este análisis de probabilidad de inasistencia y “overbooking” con una base de datos integrada a una interfaz de usuario

- resumen del resultado final obtenido

PENDIENTE

Núcleo del proyecto

Definición del proyecto y análisis de viabilidad

Descripción detallada de la empresa

MediAgenda Solutions, S.L. es una empresa líder en el sector de la salud, especializada en la administración inteligente de agendas médicas para hospitales y centros médicos en España y Latinoamérica.

Nos enfrentamos al desafío crítico de los largos tiempos de espera en la atención médica, una problemática que impacta tanto a los pacientes, quienes experimentan retrasos significativos en su atención, como a los profesionales de la salud, que deben hacer frente a agendas sobrecargadas y recursos limitados.

Un ejemplo ilustrativo de esta problemática se encuentra en el barómetro sanitario realizado en España por el por el Ministerio de Sanidad, Consumo y Bienestar Social (6) en colaboración con el Centro de Investigaciones Sociológicas. Según los datos recogidos en el informe correspondiente al año 2023, el 27.2% de los ciudadanos reportaron haber esperado "*11 días o más*" desde que solicitaron la cita hasta que fueron atendidos por el médico de familia, evidenciando así los prolongados tiempos de espera que afectan a la población.

SÓLO A QUIENES EN LOS ÚLTIMOS DOCE MESES HAN TENIDO ALGUNA CONSULTA CON UN/A MÉDICO/A DE CABECERA O DE FAMILIA DE LA SANIDAD PÚBLICA Y PASÓ MÁS DE UN DÍA HASTA QUE LES ATENDIERON PORQUE NO HABÍA CITA

ANTES
(1 o 3 en P6 y 3 en P6b)
(N=1.388)

Pregunta 6b01

¿Cuántos días?

1 día	0,1
2 días	11,4
3 días	12,7
4 días	6,8
5 días	7,3
6 días	1,8
7 días	18,5
Entre 8 y 10 días	12,4
11 y más días	27,2
No recuerda	1,7
N.C.	0,1
(N)	(1.388)
Media (días)	9,48
Desviación típica	9,47
(N)	(1.363)

Con el objetivo de abordar este problema, hemos centrado nuestros esfuerzos en el desarrollo de soluciones innovadoras que optimicen la gestión de citas médicas mediante la aplicación de tecnologías avanzadas, especialmente inteligencia artificial (IA). Nuestros avanzados algoritmos de Machine Learning analizan una amplia gama de datos, desde historiales médicos hasta patrones de comportamiento previos, para predecir con precisión la probabilidad de inasistencia de los pacientes a sus citas médicas. Esta información se utiliza para ajustar de manera inteligente las agendas médicas,

maximizando así la eficiencia de los recursos médicos disponibles y reduciendo los tiempos de espera para los pacientes.

Además de los desafíos asociados con los largos tiempos de espera en la atención médica, otro problema significativo que enfrentan tanto pacientes como profesionales de la salud es el alto porcentaje de absentismo en las citas médicas. Este fenómeno no solo genera costos económicos para los sistemas de salud, sino que también puede afectar negativamente la calidad y eficiencia de la atención médica.

Conscientes de esta problemática, en MediAgenda Solutions, S.L. hemos desarrollado una solución innovadora que aborda directamente el problema del absentismo. Nuestra plataforma utiliza algoritmos avanzados para enviar recordatorios personalizados a los pacientes días previos o incluso el mismo día de su cita médica, solicitando una confirmación de asistencia. Esto no solo ayuda a reducir el porcentaje de inasistencias, sino que también permite una mejor planificación de los recursos médicos y una atención más eficiente para aquellos pacientes que sí asisten a sus citas programadas.

La eficacia de las innovaciones tecnológicas en el ámbito médico se ha demostrado en términos de ahorro de tiempo tanto para los pacientes como para los profesionales de la salud, lo que se traduce en una reducción de costos. Estas innovaciones no solo mejoran la organización de los horarios y la reducción de tiempos de espera, sino que también permiten una actualización inmediata de la información del paciente y una mayor flexibilidad en la programación de citas.

En resumen, la inversión en tecnología aplicada a la gestión de citas médicas es crucial para optimizar los procesos, aumentar la satisfacción de los usuarios y avanzar hacia sistemas de salud más accesibles y sostenibles.

La empresa tiene su sede principal en Barcelona, España, desde donde coordina sus operaciones y desarrollo tecnológico. Sin embargo, su alcance abarca a nivel nacional e internacional, con el objetivo de ofrecer soluciones a los problemas de tiempos de espera en los servicios de salud en países de habla hispana, especialmente en España y Latinoamérica.

Análisis interno y externo

Con el propósito de obtener una visión detallada y completa de la posición y perspectivas de MediAgenda Solutions, S.L. en el sector de la salud, se realizará un análisis tanto de los factores internos como externos que influyen en la empresa.

Análisis Interno

Innovación Tecnológica: MediAgenda Solutions, S.L. se destaca por fomentar una cultura de innovación y colaboración, y se mantiene actualizada en las últimas tendencias y desarrollos en el campo de la inteligencia artificial y la tecnología de la salud.

Equipo Multidisciplinario: La empresa cuenta con un equipo altamente calificado y multidisciplinario, conformado por profesionales en áreas como inteligencia artificial, medicina, gestión de proyectos, diseño de experiencia de usuario (UX/UI) y desarrollo de software con el objetivo de desarrollar y ofrecer una solución innovadora y eficaz para

abordar el problema de los largos tiempos de espera en las consultas médicas especializadas. Esta diversidad de talentos permite una visión integral en el desarrollo de sus productos y servicios.

Alianzas Estratégicas: MediAgenda Solutions, S.L. ha establecido alianzas estratégicas con instituciones médicas, centros de salud y organizaciones del sector para colaborar en la implementación y mejora continua de sus soluciones. Estas alianzas fortalecen su posición en el mercado y les permiten acceder a una base de clientes potenciales más amplia.

Análisis Externo

Demanda en Crecimiento: La creciente demanda de soluciones para reducir los tiempos de espera en los servicios de salud representa una oportunidad clave para MediAgenda Solutions, S.L. La necesidad de optimizar la gestión de citas médicas es un problema extendido en el sector, tanto en España como en Latinoamérica, lo que brinda un mercado potencialmente amplio para sus servicios.

Competencia: Aunque MediAgenda Solutions, S.L. es líder en la integración de inteligencia artificial en la gestión de citas médicas, enfrenta competencia de otras empresas que ofrecen soluciones similares o alternativas tradicionales. La capacidad de innovación y diferenciación será crucial para mantener su posición en el mercado.

Regulaciones y Normativas: La empresa opera en un sector altamente regulado, sujeto a normativas específicas en materia de protección de datos, seguridad y calidad de servicios de salud. Cumplir con estas regulaciones es fundamental para ganar la confianza de clientes y usuarios finales, así como para mantener la reputación y credibilidad de la empresa.

Explicación detallada del proyecto

Para la ejecución de este proyecto se plantean 4 fases o retos que nos ayuden a conseguir el objetivo planteado:

- **Fase 1: Realizar un modelo de IA capaz de predecir la Asistencia o No Asistencia de los pacientes a las citas médicas.** Esta fase es esencial ya que sienta las bases para el desarrollo de soluciones efectivas en las etapas posteriores del proyecto.

- **Obtención y preparación de datos**

- Obtener una base de datos que contenga información sobre las citas médicas programadas en un servicio de salud, así como datos relacionados con los usuarios que las solicitan. Es fundamental que esta base de datos incluya una etiqueta que indique si el paciente asistió o no a su cita médica, ya que esta información es esencial para el entrenamiento y la evaluación del modelo de inteligencia artificial.
- Realizar un exhaustivo Análisis Exploratorio de los Datos (EDA, del inglés “Exploratory Data Analysis”). Esta técnica permite

comprender la estructura, la distribución y las relaciones dentro del conjunto de datos. Durante este proceso, se identifican y preparan las variables predictoras relevantes para la predicción de inasistencia a las citas médicas. Esto implicará la limpieza de datos para tratar inconsistencias o valores faltantes, así como explorar la base de datos en busca de relaciones, tendencias o anomalías. Este paso es crucial para garantizar la calidad y la confiabilidad de los datos utilizados en el entrenamiento del modelo.

○ **Feature Engineering**

El proceso de Feature Engineering desempeña un papel fundamental en la preparación de los datos del dataset para la construcción de modelos de inteligencia artificial. Durante esta etapa, transformamos los datos brutos del dataset en características significativas que permiten a los modelos aprender patrones y realizar predicciones precisas. A continuación, se detallan los pasos clave a realizar:

- **Extracción de características.** Se extraen características relevantes del dataset original, como la información demográfica de los pacientes (edad, género, ubicación) y datos relacionados con el historial médico y las visitas anteriores. Estas características son cruciales para predecir la asistencia o no asistencia a las citas médicas.
- **Transformación de tipos de datos.** Se realiza una transformación exhaustiva de los tipos de datos presentes en el conjunto de datos porque los modelos de inteligencia artificial requieren datos de un formato específico para su correcto funcionamiento. La transformación de los tipos de datos abarca varias acciones, como cambiar el formato de fechas de tipo de dato “date”. Además, se modifica el tipo de dato de la etiqueta No-show para simplificar su interpretación y su uso como variable objetivo en el modelo. Al transformarla a tipo de dato booleano, se codifica la asistencia del paciente de manera binaria (0 para asistencia, 1 para no asistencia), lo que facilita la aplicación de algoritmos de clasificación.
- **Creación de nuevas características.** Además de las características existentes, se crean nuevas características combinando variables existentes o incorporando información adicional de otras fuentes de datos que pueden tener un impacto significativo en la predicción de la asistencia o no asistencia a las citas médicas. Estas características se basan en hipótesis fundamentadas y buscan capturar aspectos relevantes del contexto de las citas. Entre las nuevas características a considerar se encuentran:
 - Información climática: Se incorporará información sobre el clima del día de la cita, ya que se ha planteado la hipótesis de que las condiciones climáticas pueden influir en la decisión del paciente de asistir a la cita médica. Esta información puede obtenerse de fuentes meteorológicas locales o bases de datos climáticas.
 - Ubicación geográfica: Se incluirá la ciudad del lugar de la cita como una característica, dado que la ubicación geográfica podría desempeñar un papel en la asistencia a las citas médicas. Se

explorará si existen diferencias significativas en los patrones de asistencia entre distintas áreas geográficas.

- Día de la semana de la cita: Se calculará el día de la semana correspondiente a la fecha de la cita médica. Se parte de la hipótesis de que ciertos días de la semana pueden estar asociados con tasas de asistencia diferentes. Por ejemplo, es posible que los pacientes tengan más disponibilidad para asistir a citas durante días laborables específicos.
- Tiempo desde la última cita: Para aquellos pacientes que tienen más de una cita registrada en el historial médico, se calculará el tiempo transcurrido desde la última cita. Esta característica se basa en la premisa de que el intervalo de tiempo entre citas puede influir en la probabilidad de asistencia a la siguiente cita. Por ejemplo, es posible que los pacientes que han esperado un período prolongado entre citas sean menos propensos a asistir.
- Tiempo de espera entre programación y cita: Se calculará el tiempo en días entre la fecha de la cita y la fecha de programación. Se plantea la hipótesis de que el tiempo de espera puede afectar la probabilidad de asistencia a la cita médica. Por ejemplo, un tiempo de espera más corto podría estar asociado con una mayor probabilidad de asistencia, mientras que un tiempo de espera prolongado podría aumentar la posibilidad de que el paciente no asista.

Basándose en la hipótesis de que el historial de un paciente concreto puede ser un buen predictor sobre la asistencia futura de ese paciente, se crean las siguientes características:

- Total de no asistencia por paciente: Esta característica implica sumar el valor de "no-show" (1 para no asistencia, 0 para asistencia) por paciente en todo su historial de citas médicas. Es decir, se contabilizará cuántas veces un paciente ha faltado a sus citas médicas en total. Esta métrica proporcionará una visión global de la tendencia de un paciente a no asistir a sus citas, lo que puede ser útil para identificar patrones de comportamiento y tomar medidas preventivas, como recordatorios de citas.
- Total de citas por paciente: Esta característica consiste en contar el número total de citas médicas registradas para cada paciente, independientemente de si ha asistido o no. Esta métrica permitirá entender la frecuencia con la que un paciente utiliza los servicios médicos, o que puede ayudar a gestionar eficientemente los recursos médicos disponibles y optimizar la planificación de citas.
- Porcentaje acumulado de no asistencia: Se calculará dividiendo el total de no asistencias entre el total de citas para cada paciente. Este porcentaje reflejará la proporción de veces que un paciente ha faltado a sus citas médicas en relación con el total de citas programadas. Proporcionará una comprensión más completa de la propensión de un paciente a faltar a sus citas médicas, considerando su historial completo de citas. Un porcentaje alto

podría indicar una mayor probabilidad de futuras ausencias

En resumen, estas nuevas características se seleccionan cuidadosamente en función de su relevancia teórica y se integran en el conjunto de datos para su posterior análisis y modelado. Su inclusión se justifica en el objetivo de mejorar la capacidad predictiva del modelo de inteligencia artificial y proporcionar insights adicionales sobre los factores que influyen en la asistencia a las citas médicas.

- ***Selección de características.*** Se identifican las características más relevantes para el problema en cuestión mediante la evaluación de su correlación con la variable objetivo (No-show). Esto implica seleccionar aquellas características que tengan un mayor impacto en la predicción, considerando la desigualdad de distribución de clases y la presencia de pacientes únicos y su historial.
- ***Manejo de datos faltantes.*** Se implementan estrategias para manejar los datos faltantes, como la imputación de valores medios o la eliminación de registros incompletos, asegurando la integridad de los resultados del modelo.
- ***Estandarización o Normalización de las características.*** Se estandarizan o normalizan las características para asegurar que todas estén en la misma escala, mejorando así la convergencia del modelo y facilitando la interpretación de los coeficientes.
- ***Reducción de la Dimensionalidad.*** Dado que el dataset original contiene un número considerable de características, se aplicarán técnicas como PCA o t-SNE para reducir su dimensionalidad. El objetivo es conservar al menos el 95% de la varianza total de los datos. En este proceso, todas las variables del dataset original, incluidas las características creadas durante el proceso de Feature Engineering, serán consideradas inicialmente. Sin embargo, la técnica de reducción de dimensionalidad seleccionará automáticamente aquellas variables que sean más relevantes para predecir el no-show, dejando únicamente aquellas que contribuyan significativamente a explicar la varianza en los datos relacionada con la variable objetivo.
- ***Manejo de Variables Categóricas.*** Las variables categóricas, como el género, la ciudad del lugar de la cita y el grupo de edad, deben ser transformadas en un formato numérico adecuado para su procesamiento por parte de los modelos de inteligencia artificial. Para lograr esto, se aplicará One-Hot Encoding (creando columnas binarias para cada categoría) para variables categóricas con múltiples categorías no ordenadas, como la ciudad del lugar de la cita. Por otro lado, se aplicará Label Encoding (asignando enteros únicos a cada categoría) para variables categóricas con categorías ordenadas, como el género. Esto asegurará que todas las características del dataset estén en un formato numérico y listas para su uso en el modelo.
- ***Balanceo de Clases.*** Dado que el dataset original presenta un desbalance significativo entre las clases, donde aproximadamente el 79.81% corresponde a pacientes que asisten a sus citas y el 20.19% a pacientes que no asisten, es fundamental aplicar técnicas de balanceo de clases para

evitar sesgos en los modelos de inteligencia artificial. Estos sesgos podrían llevar a una clasificación errónea y una falta de generalización del modelo. Para abordar este desbalance, se recurrirá a técnicas como Oversampling, Undersampling, SMOTE (Synthetic Minority Over-sampling Technique) o SMOTE-ENN (SMOTE-Edited Nearest Neighbors). Estas técnicas generan datos sintéticos adicionales para la clase minoritaria o eliminan muestras de la clase mayoritaria, con el objetivo de equilibrar el número de muestras en cada clase. Al balancear las clases de esta manera, se garantiza que el modelo aprenda de manera equitativa de todas las clases y pueda generalizar efectivamente a datos no vistos. En resumen, el balanceo de clases es crucial para mitigar el impacto del desbalance en la precisión del modelo y fomentar una clasificación justa y precisa.

- **Entrenar diferentes modelos de Machine Learning.** Dado que no se dispone de un conjunto de datos de prueba separado, se procede a designar una parte del conjunto de datos original como conjunto de prueba. Para asegurar la representatividad de este conjunto, se decide construirlo sin la etiqueta de asistencia ("No-show") y manteniendo la estratificación del conjunto original. Además, se considera la posibilidad de asegurar que el conjunto de prueba incluya solo citas de pacientes únicos, incluyendo solo la última cita para aquellos pacientes con múltiples citas.

Durante el entrenamiento de los modelos de Machine Learning, se considerarán las siguientes hipótesis:

- Se evaluará la hipótesis de crear conjuntos de datos segmentados por patología, aprovechando la información variada sobre los pacientes, como hipertensión, diabetes o alcoholismo. Esta segmentación permitiría construir modelos específicos para cada grupo de pacientes, mejorando así la capacidad predictiva. Por ejemplo, se evalúa la idea de un modelo específico para pacientes con historial de alcoholismo, ya que este factor puede influir significativamente en la asistencia a citas médicas.
- Otra hipótesis a considerar es la correlación significativa en pacientes con recurrencia, ya que se ha observado que en el dataset original un 60.87% de los pacientes se corresponde con aquellos que han tenido su primera cita. Por tanto, se plantea dividir el conjunto de datos entre datos de pacientes con recurrencia y sin recurrencia, con el fin de estudiar cada grupo por separado. Este enfoque podría revelar patrones específicos que no serían evidentes al analizar el conjunto de datos completo.
- Además, se contempla la posibilidad de realizar predicciones parciales o establecer pesos significativos para determinados perfiles de pacientes, como aquellos con problemas graves de conducta, como el alcoholismo. Esto implicaría estudiar y modelar exclusivamente los datos de este perfil para mejorar la precisión de la predicción.

Se utilizará un subconjunto más pequeño del conjunto de datos de entrenamiento como conjunto de validación para ajustar y seleccionar los mejores hiperparámetros de cada modelo. Se evaluarán varios algoritmos

de aprendizaje automático, abarcando modelos clásicos como la regresión logística y árboles de decisión, así como técnicas más avanzadas como las máquinas de vectores de soporte (SVM) y redes neuronales. Esta etapa es esencial para optimizar el rendimiento de los modelos antes de evaluarlos en el conjunto de datos de prueba.

Durante el proceso de entrenamiento y evaluación, se compararán y evaluarán las capacidades predictivas de cada modelo utilizando métricas como precisión, recall y F1-score. Esto permitirá determinar qué modelos son más adecuados para el problema en cuestión y seleccionar el modelo final o combinación de modelos que mejor se adapte a las necesidades, con la capacidad de predecir con precisión la asistencia o no asistencia a las citas médicas.

Optimizando la precisión minimizamos el número de pacientes que predecimos no asistirán, pero finalmente sí lo hacen, generándonos overbooking en la Fase 2 de nuestro proyecto. Es decir, minimizamos el overbooking real generado.

Optimizando el recall minimizamos el número de pacientes que predecimos que asistirán, pero finalmente no los hacen, generándonos huecos libres en las agendas médicas. Es decir, minimizamos las pérdidas por inactividad en la Fase 2 de nuestro proyecto.

● **Fase 2:** Una vez finalizada la fase 1, y habiendo obtenido una probabilidad de asistencia a las citas médicas, la fase de **implementar un sistema de overbooking en la programación de citas** para maximizar el uso de los recursos de los centros hospitalarios sin desbordar los costos provocados por la propia implementación del sistema con overbooking no es sencilla ni determinista.

Son varios los factores que complican la resolución real del problema, los cuales se exponen a continuación.

- **Incertidumbre en las predicciones de asistencia.** Los modelos de predicción están lejos de ser perfectos. Aunque los resultados son mejores que una mera asignación de probabilidad media de asistencia o una predicción al azar, las métricas de precisión y sensibilidad siguen proporcionando un grado de incertidumbre tal que será automáticamente trasladado a nuestros modelos de optimización.
- **Variedad e incertidumbre en los factores ambientales.** Para empezar a tratar este tema, el número de factores ambientales que se pueden aplicar al sistema son muy numerosos, prácticamente diferentes para cada centro hospitalario o de salud. Adicional, muchos de ellos añaden un gran valor de incertidumbre que ni siquiera está estudiado por modelos predictivos. Así pues, podemos distinguir entre factores más certeros (número de servicios entregados, número de doctores, número de citas por sesión, prioridad en la atención) y más inciertos (impuntualidades, tiempos de atención, nivel de interrupción de los doctores, llegadas espontáneas).

- **Variedad en las reglas de cita y en reglas de secuencia.** Las reglas de cita definen las restricciones que imponemos al modelo para limitar las asignaciones (número de pacientes que se puede asignar a un único slot, duración de los slots, etc.), y las reglas de secuencia definen el orden con los que los pacientes son asignados a los slots en función de una determinada clasificación (prioritarios, primerizos, recurrentes, etc.). Sólo Cayirli, Veral et Rosen (7) ya realizaron un estudio probando 6 diferentes reglas de secuencia con 7 reglas de cita, lo que hace un total de 42 Sistemas de Asignación (AS, del inglés) diferentes. Para hacernos otra idea de la magnitud y complejidad del problema, Ho y Lau (8) comparaban 9 diferentes sistemas de citas en 27 diferentes escenarios clínicos caracterizados por 3 factores ambientales (probabilidad de ausencias, variación de tiempos de atención y número de pacientes por sesión).
- **Definición del Coste.** Para minimizar la Función de Coste que tenemos que resolver para hacer la mejor programación de citas médicas primero tenemos que definir cuál es el coste y cómo calcularlo. Es mucha la literatura al respecto, Chew (9) divide la función en 3 tipologías de coste distintas: el tiempo de espera del paciente en consulta, el tiempo de inactividad de toda la infraestructura médica por falta de paciente, y el tiempo extra dedicado por la infraestructura médica a atender los pacientes después de una jornada laboral normal. Almaktoom (10) calcula el coste respecto a los costos de realizar una consulta, los de operación de la clínica y los de overbooking, referenciándolos al número de No-Shows y de Overbookings. Sin embargo, en Valenzuela-Núñez et al. (11) no consideran minimizar un coste, sino maximizar la utilidad que genera el centro sanitario por cada cita atendida, menos una penalidad por sobrecarga de overbooking, directamente proporcional a los casos de overbooking registrados. Y Lawley and Muthuraman (12) usaron un modelo estocástico para minimizar el número de pacientes desbordados de una cita a la siguiente.

En nuestro caso, en aras de simplificar la resolución del problema, vamos a minimizar las variables del modelo:

- **Factores ambientales:**
 - Predicciones de Asistencia. Extraídas de la Fase 1.
 - Número de slots por sesión = 16 (consideramos consultas de media hora en una jornada normal de 8 horas). **REVISABLE**
 - El resto de los factores los anulamos directamente: presuponemos que no hay impuntualidad, que los tiempos de atención son los mismos que las duraciones de los slots, que no hay llegada espontánea de pacientes, un único servicio entregado, etc.
 - Asistencias reales.
- **Reglas de citas:**
 - Dentro de esta variable están los parámetros que deseamos que nuestro modelo aprenda para minimizar el coste:

- Número de pacientes a solapar en cada slot (grado de overbooking). Es decir, tantos parámetros distintos como slots definamos en una sesión.
- Duración de slot = 30 minutos. **REVISABLE**

○ **Reglas de secuencia:**

- Utilizaremos una secuencia FCFA (First Call First Appointment), también conocida directamente como Regla de No Secuencia, pues simplemente se asignará al paciente en el mejor slot disponible calculado mediante el modelo entrenado de overbooking.

○ **Función de Coste:**

- Utilizaremos una fórmula parecida a la usada por Chew (9), dividiendo y calculando la función de coste en 3 términos independientes:
 - T_P = Tiempo de espera del paciente. Calculado como unidad temporal que debe esperar un paciente a ser atendido por encontrarse el slot sobrecargado. Si tiene que esperar será igual a 1, y así sucesivamente. (unidad temporal = 30 min = 1 slot).
 - T_D = Tiempo extra a realizar por el doctor. Calculado como unidad temporal que tiene que trabajar fuera de la sesión o jornada laboral normal de 8 horas. Si no tiene que trabajar ningún slot será igual a 0, si tiene que trabajar un slot será igual a 1, y así sucesivamente. (unidad temporal = 30 min = 1 slot).
 - T_I = Tiempo de inactividad en la clínica. Calculado como unidad temporal sin estar atendiendo a ningún paciente. Si el doctor está siempre ocupado será igual a 0, si queda un slot vacío será igual a 1, y así sucesivamente. (unidad temporal = 30 min = 1 slot).
- A cada término independiente le daremos pesos distintos en función de cuál sea el coste que más queremos penalizar, en el entendido de que hay valores muy subjetivos en el cálculo de dichos pesos, como pueden ser el coste que tiene para un paciente la hora de espera, o para un doctor la hora extra de trabajo extra.

Entrenaremos el modelo para que calcule los mejores parámetros (grados de overbooking). El punto crítico del modelo es calcular inicialmente el escenario ideal de mínimo coste con los datos de las asistencias reales conocidas, para luego ajustarlo con las predicciones de asistencia obtenidas de la Fase 1 de forma que se minimicen las diferencias de coste (no el coste propiamente dicho, pues éste ya se ha minimizado en la 1ª pasada).

Posteriormente calcularemos las pérdidas obtenidas en un escenario determinado con asistencias reales, ejecutando una programación de agenda médica utilizando únicamente sus correspondientes predicciones de asistencia.

En ambos casos, entrenamiento y prueba, nos hemos de asegurar de que haya citas suficientes como para rellenar todos los slots de varias sesiones (días) y generar overbookings. En caso necesario, se generarán muestras sintéticas.

Finalmente compararemos los resultados obtenidos en el modelo entrenado con otros modelos sin entrenar, más rígidos en las restricciones o reglas de citas, así como un modelo en donde no hagamos overbooking alguno.

● **Fase 3: Creación de un Asistente Virtual basado en el Procesamiento de Lenguaje Natural (NLP)** que permita el contacto entre Centros de Salud y Pacientes para la gestión de citas y recordatorios.

En esta fase, se desarrollará un Asistente Virtual que permitirá la comunicación entre los

Centros de Salud y pacientes para la gestión de citas y recordatorios. El objetivo principal es mejorar la experiencia del usuario y optimizar el proceso de programación y seguimiento de citas médicas.

El Asistente Virtual se basará en técnicas de Procesamiento de Lenguaje Natural (NLP) para comprender las consultas y solicitudes de los pacientes, y proporcionar respuestas relevantes y precisas. A continuación, se detallan los pasos a seguir en esta fase:

○ **Objetivos del Asistente Virtual:**

- Gestión de citas:
 - Permitir a los pacientes solicitar, modificar y cancelar citas médicas a través de WhatsApp.
 - Proporcionar información sobre la disponibilidad de citas y los horarios de los médicos.
 - Confirmar las citas programadas y enviar recordatorios a los pacientes.
- Proporción de información relevante:
 - Brindar información general sobre los servicios y especialidades médicas disponibles en los Centros de Salud.
 - Responder a preguntas frecuentes relacionadas con el proceso de atención médica, documentación requerida y protocolos de seguridad.
 - Ofrecer orientación básica sobre síntomas comunes y medidas preventivas, siempre enfatizando la importancia de consultar a un profesional de la salud.
- Envío de recordatorios:
 - Enviar recordatorios automatizados a los pacientes sobre sus próximas citas programadas.
 - Notificar a los pacientes sobre cualquier cambio o actualización relevante en relación con sus citas.
 - Proporcionar instrucciones y recomendaciones previas a la cita, como la preparación necesaria o la documentación requerida.

○ **Alcance del Asistente Virtual:**

- El Asistente Virtual podrá manejar consultas relacionadas con la programación, modificación y cancelación de citas médicas.
 - Responderá a preguntas generales sobre los servicios médicos, horarios de atención y disponibilidad de especialistas.
 - Proporcionará información básica sobre síntomas comunes y medidas preventivas, siempre remitiendo a los pacientes a un profesional de la salud para un diagnóstico y tratamiento adecuados.
 - Ofrecerá asistencia en la navegación del proceso de atención médica, como los pasos a seguir para programar una cita o los documentos necesarios.
 - Enviará recordatorios y notificaciones relacionados con las citas programadas.
 - Recopilará información básica de los pacientes, como nombre, fecha de nacimiento y motivo de la consulta, para facilitar el proceso de programación de citas.

○ **Generación de Preguntas Relevantes (Prompt Engineering):**

- Se creará un conjunto de preguntas clave para obtener la información necesaria de los pacientes, como el motivo de la consulta, preferencias de fecha y hora, y datos personales básicos.
- Se desarrollarán prompts específicos para cada etapa del proceso de programación de citas, desde la solicitud inicial hasta la confirmación final.
- Se utilizarán técnicas de Prompt Engineering, como la personalización y la contextualización, para generar preguntas claras, concisas y relevantes.
- Se implementarán mecanismos de manejo de errores y aclaraciones para garantizar la comprensión adecuada de las respuestas de los pacientes.

○ **Optimización de los Flujos de Conversación y la Comprensión de Texto:**

- Se diseñarán flujos de conversación intuitivos y lógicos para guiar a los pacientes a través del proceso de programación de citas de manera eficiente.
- Se utilizarán técnicas de NLP, como la tokenización y el análisis de entidades, para extraer información clave de las consultas de los pacientes.
- Se implementarán algoritmos de detección de intenciones para identificar el propósito principal de cada consulta y proporcionar respuestas adecuadas.
- Se aplicarán técnicas de manejo de contexto para mantener la coherencia y la relevancia en la conversación.
- Se desarrollarán mecanismos de resolución de ambigüedades y manejo de errores para garantizar una comunicación clara y precisa.

○ **Integración del Asistente en la plataforma de mensajería (WhatsApp):**

- Se utilizará la API de WhatsApp Business para integrar el Asistente Virtual en la plataforma de mensajería.
- Se desarrollarán funciones para recibir y procesar los mensajes entrantes de los pacientes.
- Se implementarán mecanismos de respuesta automatizada para proporcionar información y asistencia de manera oportuna.
- Se crearán funciones para enviar mensajes, recordatorios y notificaciones a los pacientes.
- Se manejarán los flujos de conversación y las interacciones con los pacientes de manera eficiente y escalable.

La implementación del Asistente Virtual se realizará utilizando la API de OpenAI, aprovechando sus capacidades de procesamiento de lenguaje natural. Se emplearán instrucciones de sistema para establecer el contexto y el comportamiento general del asistente. Además, se utilizarán funciones específicas para tareas como la programación de citas, la proporción de información. También se aplicarán técnicas de retrieval para acceder a información relevante de la base de datos de citas médicas.

La integración con la base de datos permitirá al Asistente Virtual acceder a información en tiempo real sobre la disponibilidad de citas, los horarios de los médicos y los detalles de las especialidades médicas. Esto garantizará que las respuestas proporcionadas a los pacientes sean precisas y actualizadas como se puede ver en la figura 1:

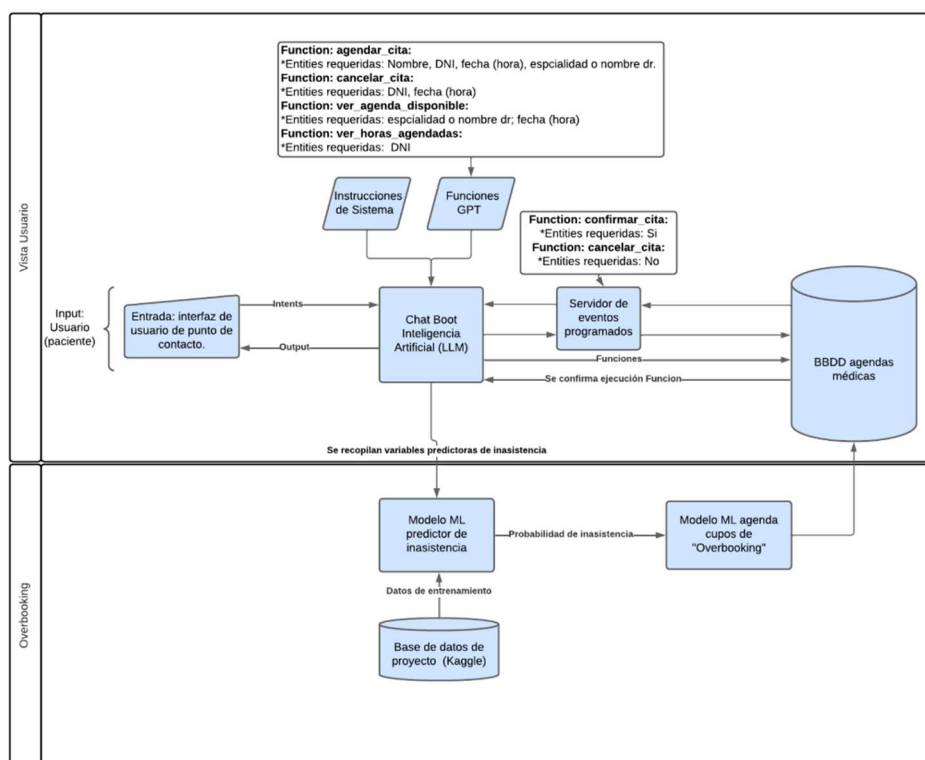


Figura 1. Diagrama de Flujo propuesta de Asistente Virtual.

- **Fase 4: Desarrollar**, a nivel teórico, **un proceso de refinamiento del sistema que integre la base de datos de la interfaz creada para gestionar las citas y los recordatorios a los pacientes** (Fase 3) con los modelos predictivo y de optimización generados en las Fases 1 y 2.

- Selección de posibles nuevas características relevantes provenientes de la aplicación.
- Adhesión de estas nuevas características predictivas al dataset definitivo elaborado en la Fase 1.
- Desarrollo fine-tuning del modelo de IA seleccionado en la Fase 1, con el objetivo de mejorar en precisión las predicciones de asistencia.
- Integrar los nuevos resultados en modelo de IA creado en la Fase 2, analizando la optimización / disminución de “coste” obtenida.
- Fusionar este análisis de probabilidad de inasistencia y “overbooking” con una base de datos integrada a una interfaz de usuario.

Fijación de los objetivos generales y específicos del proyecto

El proyecto, tal como su propio título indica, se basa en el uso de Inteligencia Artificial para la gestión de citas médicas, innovando en la confirmación y reducción de ausencias. El objetivo general es reducir los costes asociados a los tiempos de inactividad en una clínica, es decir, aquellos generados por una falta de aprovechamiento de los recursos humanos y materiales de los centros médicos, y que, indirectamente, genera otros costos adicionales al sistema de salud de difícil cuantificación (listas de espera excesivas, calidad del servicio médico, etc.).

En nuestro caso particular, teniendo el proyecto bien definido en 4 fases, dividimos este objetivo general en otros más concretos dentro de cada fase, definiendo con ellos los objetivos específicos de cada una de ellas:

- **Fase 1:**

Desarrollar un modelo de Inteligencia Artificial que, dada una base de datos con información específica de pacientes e información histórica de las visitas médicas realizadas anteriormente, sea capaz de predecir las No Asistencias a futuras citas médicas con una precisión mayor a una predicción al azar o en base a una media probabilística de asistencias.

El objetivo específico, para conseguir posteriormente hacer una buena optimización del modelo de programación de citas médicas, es obtener la mejor F1-Score posible, de forma que tengamos la mejor combinación entre una buena precisión: tener pocos False Positives (pacientes que predecimos que no asistirán y en realidad sí lo hacen), es decir, disminuir los costos de espera de los pacientes y de trabajo extra por parte de los doctores; y una buena recall: tener pocos False Negatives (pacientes que predecimos que asistirán y en realidad no lo hacen), es decir, disminuir el coste de inactividad de la clínica.

- **Fase 2:**

Conseguir optimizar un modelo entrenado con IA para que haga una programación de citas de forma inteligente, utilizando técnica de overbooking,

minimizando el costo producido tanto por la aparición de huecos en la agenda (tiempo inactividad clínica) como por la aparición de efectos secundarios generados con el overbooking (tiempo espera del paciente y tiempo extra del doctor).

Específicamente, este modelo deberá generar menos coste que un modelo donde las citas se asignen directamente, sin considerar las probabilidades de no asistencia, es decir, sin usar técnicas de overbooking. Adicionalmente, revisaremos que el modelo generado con IA también genere menor coste que otros modelos tradicionales de programación de agendas basados en reglas de citas y de secuencia.

- **Fase 3:**

Generar un Asistente Virtual que establezca un canal de comunicación robusto, amigable y fácil de usar, capaz de comunicarse eficazmente con los pacientes, asignar citas de forma precisa, enviar recordatorios, y conseguir confirmaciones de asistencia.

Dentro de la Gestión de Citas es importante cumplir con los siguientes objetivos más específicos:

- Proporcionar información sobre la disponibilidad de citas y los horarios de los médicos.
- Permitir a los pacientes solicitar, modificar y cancelar citas médicas a través de WhatsApp.
- Enviar recordatorios automatizados a los pacientes sobre sus próximas citas programadas, confirmando las citas programadas.
- Notificar a los pacientes sobre cualquier cambio o actualización relevante en relación con sus citas.
- Proporcionar instrucciones y recomendaciones previas a la cita, como la preparación necesaria o la documentación requerida.

También se considerará, a nivel teórico, la capacidad del Asistente Virtual para ser entrenado con los siguientes objetivos secundarios:

- Brindar información general sobre los servicios y especialidades médicas disponibles en los Centros de Salud.
- Responder a preguntas frecuentes relacionadas con el proceso de atención médica, documentación requerida y protocolos de seguridad.
- Ofrecer orientación básica sobre síntomas comunes y medidas preventivas, siempre enfatizando la importancia de consultar a un profesional de la salud.

- **Fase 4:**

Desarrollada muy a nivel teórico, el objetivo de esta fase es desarrollar un hipotético proceso de mejora continua, donde el modelo predictivo y la información extraída por el Asistente Virtual se retroalimenten periódicamente para mejorar los resultados predictivos del modelo, para que, de forma análoga, dichas mejoras se trasladen al modelo de optimización de agenda médica, reduciéndose los costes asociados a inasistencias.

Planificación

Estimación de recursos económicos

Para poder realizar esta estimación se espera a las decisiones tomadas por las directivas, posterior al estudio de la prefactibilidad.

Estimación de recursos materiales

Para estimar los recursos materiales del asistente virtual se deben tener en cuenta los siguientes recursos:

- Un servidor o infraestructura en la nube, para alojar y ejecutar el programa del asistente virtual que se decida usar; debe tener capacidad suficiente para procesar y almacenar la data. Y una amplia capacidad de procesar un alto volumen de citas.
- Dispositivos de acceso como computadores - terminales, tablets o teléfonos inteligentes.
- Línea de acceso a Internet, ¿canal dedicado?
- Software licenciado para el programa - Incluye sistema operativo, Python, Visual Studio, u otras soluciones más robustas, que se decida adoptar; librerías y herramientas para los procesamientoos.
- Otro servidor o infraestructura en la nube, para alojar y ejecutar los programas de prueba o nuevos desarrollos.
- Un sistema de gestión para almacenamiento de las bases de datos, y de copias de respaldo.
- El espacio físico para alojar los servidores y las personas responsables del proyecto.

Estimación de recursos humanos.

Las personas que harán parte del proyecto de asistente virtual se pueden dividir en dos momentos, planeación, desarrollo e implementación del proyecto y personal de funcionamiento y mantenimiento del asistente virtual.

- La primera parte debe contar con al menos un desarrollador de software.

Para todo el proyecto se necesitan personas de coordinación y soporte, como son las que siguen:

- Gerente del proyecto.
- Especialistas de seguridad de datos.
- Personal de soporte para hardware y software.
- Al menos un experto en IA, inicialmente puede ser medio tiempo.

De otro lado se debe tener en cuenta los implicados directamente con el proyecto que deben tener representación en todo el desarrollo del proyecto, al menos un médico, personal administrativo, personal de atención al cliente.

Estimación de los recursos tiempo

Se ha estimado que para el estudio de prefactibilidad un mes.

En cuanto hace referencia a la planeación, desarrollo e implementación del proyecto: un mes y medio.

Para la parte de pruebas y ajustes un mes.

Lo anterior está condicionado a la cantidad de personas que puedan contratarse para este proyecto.

Elaboración del cronograma del proyecto

Se ha estimado el siguiente cronograma del proyecto:

- | | |
|--|---------------|
| ○ Fase de estudio de prefactibilidad: | Día 00 - 30 |
| ○ Presentación ante las directivas: | Día 31 - 45 |
| ○ Planeación, desarrollo e implementación: | Día 46 - 60 |
| ○ Validación, pruebas y ajustes: | Días 61 - 90 |
| ○ Despliegue e implementación: | Días 91 - 120 |
| ○ Salida a producción: | Día 121 |

Definición del alcance del proyecto

Una vez determinada la necesidad de realizar el presente proyecto por la institución se ha procedido a realizar las actividades descritas. Y continuamos describiendo el alcance del asistente virtual.

Recolección de datos, se obtuvo una base de datos sobre la que se realizará el desarrollo del presente asistente virtual. En el que se están desarrollando actividades de limpieza de datos, errores y manejo de datos atípico.

Integración con el sistema de gestión de citas, que se va a ir realizando a medida que avanza el proyecto, que incluye la implementación de una interfaz de programación de aplicaciones (API - Application Programming Interfaces).

Desarrollo del modelo del asistente virtual, núcleo del presente trabajo que se ha venido diseñando y adaptando desde el inicio de este proyecto.

Validación y pruebas, a medida que se va avanzando con el proyecto, se evaluará el rendimiento del asistente virtual, usando datos de prueba y validación; pruebas de carga para verificar la escalabilidad y rendimiento, bajo diferentes condiciones de carga.

Despliegue e implementación, se realizará capacitación al personal involucrado en la asignación de citas, asistentes de atención al cliente, personas del área administrativa, médicos y personal de enfermería. Se realizará supervisión y soporte directo por un periodo de tiempo por definir.

Entregables del proyecto, se plantea realizar un documento de los requerimientos del sistema de asistencia virtual, el conjunto de los datos preprocesados, el modelo del asistente virtual entrenado y desplegado. Informe de las pruebas y validación, sistema de asistente virtual implementado y en funcionamiento, así como la documentación de soporte y capacitación.

Bibliografía

- (1) Ministerio de Sanidad, Consumo y Bienestar Social. Informe Anual del Sistema Nacional de Salud 2022. [Internet]. Madrid: Ministerio de Sanidad, Consumo y Bienestar Social; 2022 [citado 30 de marzo de 2024]. Disponible en: https://www.sanidad.gob.es/estadEstudios/estadisticas/sisInfSanSNS/tablasEstadisticas/InfAnualSNS2022/INFORME_ANUAL_2022.pdf
- (2) Ministerio de Salud de Chile. Ordinario N° 331 - Glosa 06 IV Trimestre 2022. [Internet]. Santiago: MINSAL; 2022 [citado el 30 de marzo de 2024]. Disponible en: <https://www.minsal.cl/wp-content/uploads/2021/05/Ord.-331-Glosa-06-IV-Trim-2022.pdf>
- (3) Danke K, Delteil P, Figueroa R, Infante A, Orellana R, Palavicino P. Estudio de brechas de médicos y odontólogos generales y especialistas en el sector público de salud período 2020-2030. Disponible en: <https://www.minsal.cl/wp-content/uploads/2021/10/INFORME-ESTUDIO-BRECHAS-MEDICOS-Y-ODONTOLOGOS.pdf>
- (4) Rico JP. Inasistencia horas médicas: La oportunidad para implementar un modelo de atención digital [Internet]. Tierramarillano. 2023 Jun 27 [[citado el 30 de marzo de 2024]. Disponible en: <https://tierramarillano.cl/2023/06/27/inasistencia-horas-medicas-la-oportunidad-para-implementar-un-modelo-de-atencion-digital/>
- (5) <https://tekiosmag.com/2024/01/26/como-el-uso-de-tecnologia-ha-permitido-disminuir-el-no-show-de-pacientes-a-sus-citas-medicas/>
- (6) Barómetro Sanitario 2023 (Tercera oleada). Distribuciones marginales. Estudio nº3426. Octubre 2023. <https://www.cis.es/documents/d/cis/es3426mar-pdf>
- (7) T. Cayirli, E. Veral and H. Rosen. Designing appointment scheduling systems for ambulatory care services. Accepted: October 2005. Disponible en: https://www.researchgate.net/publication/7167217_Designing_Appointment_Scheduling_Systems_for_Ambulatory_Care_Services
- (8) Chrwan-Jyh H. and Hon-Shiang L. Minimizing Total Cost in Scheduling Outpatient Appointments. December 1992. Disponible en: https://www.researchgate.net/publication/227446365_Minimizing_Total_Cost_in_Scheduling_Outpatient_Appointments
- (9) Song Foh Chew. Outpatient Appointment Scheduling with Variable Interappointment Times. Accepted: July 2011. Disponible en: <https://www.hindawi.com/journals/mse/2011/909463/>

- (10) A.T. Almaktoom. Health care overbooking cost minimization model. July 2023. Disponible en: <https://europepmc.org/article/med/37560686#sec3>
- (11) C. Valenzuela-Núñez, F. Troncoso and G. Latorre-Núñez. Medical Appointment Scheduling Problem meets Machine Learning and Optimization. Ene 2024. Disponible en: <https://ieeexplore.ieee.org/document/10380593>
- (12) M. Lawley and K. Muthuraman. Stochastic Overbooking Model for Outpatient Clinical Scheduling with No-Shows. July 2008. Disponible en: <https://www.tandfonline.com/doi/full/10.1080/07408170802165823?needAccess=true>

BORRADOR

Bibliografía

1. **Ministerio de Sanidad, Consumo y Bienestar Social, Gobierno de España.** *Informe Anual del Sistema Nacional de Salud.* 2022.
2. **Ministerio de Salud, Gobierno de Chile.** *Lista de Espera No Ges y Garantías de Oportunidad GES retrasadas. Glosa 06. IV Trimestre.* 2022.
3. **Danke, Karen, y otros.** *Estudio de brechas de médicos y odontólogos generales y especialistas en el sector público de salud. Período 2020-2030.* 2020.
4. **Rico, Juan Pablo.** Inasistencia horas médicas: La oportunidad para implementar un modelo de atención digital. [En línea] 2023.
<https://tierramarillano.cl/2023/06/27/inasistencia-horas-medicas-la-oportunidad-para-implementar-un-modelo-de-atencion-digital/>.
5. **Cruz, Martín.** Cómo el uso de tecnología ha permitido disminuir el “no show” de pacientes a sus citas médicas. [En línea] 2024. <https://tekiosmag.com/2024/01/26/como-el-uso-de-tecnologia-ha-permitido-disminuir-el-no-show-de-pacientes-a-sus-citas-medicas/>.
6. **Centro Investigaciones Sociológicas (CIS).** *Estudio n°3426. Barómetro Sanitario 2023 (tercera oleada).* 2023.
7. *Designing Appointment Scheduling Systems for Ambulatory Care Services.* **Cayirli, Tugba, Veral, Emre A. y Rosen, Harry.** s.l. : Health Care Manage Sci, 2005, Vol. 9.
8. *Minimizing Total Cost in Scheduling Outpatient Appointments.* **Chrwan-Jyh, Ho y Hon-Shiang, Lau.** 12, s.l. : Management Science, 1992, Vol. 38.
9. *Outpatient Appointment Scheduling with.* **Chew, Song Foh.** s.l. : Hindawi Publishing Corporation, 2011, Vol. 2011.
10. *Health care overbooking cost minimization model.* **Almaktoom, Abdulaziz T.** s.l. : Heliyon, 2023, Vol. 9.
11. *Smart Medical Appointment Scheduling: Optimization, Machine Learning, and Overbooking to Enhance Resource Utilization.* **Valenzuela-Núñez, Catalina, Latorre-Núñez, Guillermo y Troncoso Espinosa, Fredy.** s.l. : IEEE Acces, 2024, Vol. 12.
12. *A stochastic overbooking model for outpatient clinical scheduling with no-shows.* **Muthuraman, Kumar y Lawley, Mark.** s.l. : IIE Transactions, 2008, Vol. 40.