



Universidad de Jaén

**Escuela Politécnica Superior
de Jaén**

PREDICCIÓN DE LAS PATOLOGÍAS MÁS COMUNES Y EMERGENCIAS SANITARIAS BASADA EN VARIABLES ESPACIALES MEDIANTE EL USO DE LA MINERÍA DE DATOS

Autor: Juan Carlos Arias Ortas

Directores de la tesis:

M. Isabel Ramos Galán

Departamento: Ingeniería Cartográfica, Geodésica y Fotogrametría

Juan José Cubillas Mercado

**Departamento: Tecnologías de la información y las Comunicaciones aplicadas a
la Educación (Universidad Internacional de La Rioja)**

Fecha: 03/05/2024

**ISBN:
Licencia CC**

A mi mujer y mis hijos: Inmaculada, Marina y Juan Carlos.

AGRADECIMIENTOS

Sin duda la persona a la que agradezco de forma infinita su ayuda para poder realizar esta tesis y sin la cual hubiera sido imposible realizarla es a mi mujer Inmaculada. Esta ayuda no sólo se ha limitado al apoyo diario necesario al tener que dedicarle tiempo a este trabajo, sino también en el diseño y construcción de presentaciones, artículos y de esta misma tesis. Su apoyo y ánimo constante durante todo estos años ha sido indispensable para poder llegar a este momento.

Agradezco a mi hija Marina y a mi hijo Juan Carlos el apoyo incondicional durante todo este periodo de investigación.

A Marina todo su conocimiento y ayuda técnica en cualquiera de las fases por las que he ido pasando, presentaciones, artículos, traducciones, etc... Estando siempre disponible sin importarle días, fines de semanas, ni horas...

A Juan Carlos su apoyo constante e incondicional en tantas horas de estudio que hemos compartido. Le agradezco también los conocimientos sanitarios que me ha aportado en esta tesis, que está adquiriendo recientemente y que no ha dudado en poner a mi disposición.

Agradezco muy especialmente a mi tutora y directora María Isabel Ramos el apoyo, ayuda y dedicación que me ha dado durante todos estos años. Su increíble capacidad de trabajo, fuerza de voluntad y constancia ha sido fundamental para ir avanzando hacia nuestra meta. Su impulso ha sido vital en muchos momentos de esta investigación.

A mi director Juan José Cubillas le agradezco enormemente todos sus conocimientos técnicos y apoyo diario. Sus grandes conocimientos han hecho posible encontrar soluciones rápidas y eficaces a todos los problemas que nos hemos encontrado a lo largo de toda la investigación. Le agradezco la paciencia al atenderme y dar respuesta a todas mis dudas en cualquier día.

Finalmente, quiero agradecer a Francisco Feito la oportunidad y confianza que me dio hace años de emprender esta aventura. Sin ese apoyo nada de esto existiría, sus consejos y ánimos hicieron que esas ideas iniciales fueran tomando forma y se pudieran materializar.

Resumen

El sector sanitario es uno de los más sensibles de nuestra sociedad. La atención oportuna en el sector sanitario es esencial para la recuperación de los pacientes, y más aún en caso de emergencia sanitaria. En estos casos, es esencial una gestión adecuada de los recursos humanos y técnicos, que son limitados y deben movilizarse de forma óptima y eficaz. Se considera que la aplicación de técnicas específicas y detalladas de creación y diseño de bases de datos puede mejorar la calidad de la atención a los pacientes.

Primera Parte: Creación de un sistema de base de datos multifuente.

Crear un sistema de base de datos que almacene de forma normalizada y conjunta toda la información que pueda estar relacionada directa o indirectamente con la aparición y gestión de una emergencia sanitaria. Se ha desarrollado una metodología para gestionar e integrar en una base de datos centralizada datos sanitarios, sociales, económicos, medioambientales y geoespaciales relacionados con cada una de las demandas de emergencia.

Segunda Parte: Diseño capaz de predecir el número recursos sanitarios necesarios para atender a las urgencias y emergencias sanitarias de una zona determinada.

Se utilizan algoritmos de Machine Learning en los que las variables de entrada son datos multifuente y el atributo objetivo es la predicción del número de demandas de

emergencias sanitarias que se producirán para una fecha seleccionada. Se integraron y relacionaron datos sanitarios, sociales, económicos, medioambientales y geoespaciales relacionados con cada una de las demandas de emergencia. Se utilizaron algoritmos de regresión lineal y no lineal: máquina de vectores soporte (SVM) con kernel lineal y modelo lineal generado (GLM), y la SVM no lineal con kernel gaussiano.

Estudio del nivel de interrelación entre todas las variables relacionadas con la atención al paciente que nos permita tomar decisiones estratégicas orientadas a la creación de modelos predictivos.

Tercera Parte: Generación de un modelo capaz de predecir el número de recursos necesarios para asistir a las demandas de patologías respiratorias en un periodo y lugar determinado.

Después de analizar los distintos tipos de demanda existentes en la atención de urgencias y emergencias, observamos que uno de los motivos más frecuentes en este sector de la sanidad son los casos de enfermedades respiratorias. Los pacientes con patologías respiratorias son uno de los grupos más numerosos y por tanto un objetivo perfecto para nuestra investigación.

Como adelanto de nuestra investigación, indicar que el análisis previo de este conjunto de datos, demostró que será necesario separar la predicción de los días laborables de la de los días festivos o fines de semana. Una posible razón de ello puede ser que los motivos que llevan a la demanda de servicios de emergencia son diferentes en ambos casos, por lo que se requiere un estudio separado.

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

INDICE GENERAL

1. INTRODUCCIÓN.....	19
1.1 ESTADO DEL ARTE	23
1.2 HIPÓTESIS Y APORTACIONES DE ESTE TRABAJO..	28
2. METODOLOGÍA.....	34
2.1 ÁREA GEOGRÁFICA DE LA INVESTIGACIÓN	34
2.2 DESCRIPCIÓN DE LA METODOLOGÍA DE TRABAJO EN EL ÁREA DE EMERGENCIAS EN ANDALUCÍA.....	39
2.3 OBJETIVOS DE LA TESIS	42
2.4 INTEGRACIÓN Y GESTIÓN DE DATOS.....	43
2.5 ALGORITMOS DE MINERÍA DE DATOS	48
2.5.1 Algoritmos de regresión lineal	49
2.5.2 Algoritmos de regresión No Lineal.....	52
3. DISEÑO DE LA INVESTIGACIÓN.....	58
3.1 LA BASE DE DATOS	58
3.1.1 Bloques de información integrados	59
3.1.1.1 Bloque de información sanitaria.....	59
3.1.1.1.1 Datos personales del paciente	59
3.1.1.1.2 Información sanitaria del paciente.....	60
3.1.1.1.3 Información cronológica de la asistencia.....	62
3.1.1.2 Información meteorológica	64
3.1.1.3 Información sociológica.....	66
3.1.2 DISTRIBUCIÓN EN CUADRANTES GEOLOCALIZADOS DE JAÉN.	66
3.1.3 Diagrama Entidad-Relacion.....	69
3.1.4 Tablas, depuración de tablas y preparación de datos.	70

3.1.4.1	Tablas principales utilizadas	70
3.1.4.2	Depuración de datos.....	82
3.1.4.2.1	Primer proceso de depuración.	84
3.1.4.2.2	Segundo proceso de depuración.....	84
3.2	MODELO PREDICTIVO DE RECURSOS SANITARIOS EN CASOS DE DIFICULTAD RESPIRATORIA O DISNEA.....	94
3.2.1	Estudio y análisis del estado actual.....	94
3.2.2	Técnicas de minería de datos utilizadas	101
4.	RESULTADOS	106
4.1	GENERACIÓN DEL MODELO PARA DÍAS LABORABLES.....	106
4.2	GENERACIÓN DE MODELOS EN FESTIVOS Y FINES DE SEMANA	111
5.	DISEÑO DEL MODELO DE DATOS	118
6.	DISCUSIÓN	124
7.	CONCLUSIONES.....	128
8.	POSIBLES NUEVOS ESTUDIOS.....	134
8.1	AMPLIACIÓN A OTRAS DEMANDAS ASISTENCIALES ...	134
8.2	RED DE TRANSPORTE URGENTE.....	134
9.	DIFUSIÓN DEL ESTUDIO	138
10.	BIBLIOGRAFÍA.....	146
	ANEXO	160

INDICE DE FIGURAS

FIGURA 1. RECURSOS GENERALES	21
FIGURA 2. SERVICIOS PROVINCIALES CES 061.	21
FIGURA 3 PROCESO GENERAL.....	22
FIGURA 4 LOCALIZACIÓN DE LA ZONA DE ESTUDIO EN JAÉN, ANDALUCÍA (ESPAÑA).....	35
FIGURA 5. PLANTILLA ASISTENCIAL POR PROVINCIAS.	37
FIGURA 6. TIEMPOS MEDIOS DE RESPUESTA EN ZONAS URBANAS EN ANDALUCÍA.	38
FIGURA 7. PACIENTES ATENDIDOS POR EQUIPOS 061 EN ANDALUCÍA.	39
FIGURA 8 CICLO COMPLETO DE GESTIÓN DE URGENCIAS Y EMERGENCIAS.	41
FIGURA 9 CONSTRUCCIÓN BASE DE DATOS RELACIONAL MULTIMODAL.	45
FIGURA 10 OBJETIVOS DE LA INVESTIGACIÓN.	46
FIGURA 11 SVM KERNEL LINEAL.....	51
FIGURA 12 SVM KERNEL RADIAL.	52
FIGURA 13 SVM KERNEL GAUSSIANO RBF, PASO DE 2D A 3D	54
FIGURA 14 COMPOSICIÓN DEFINITIVA DE LA TABLA METEOROLÓGICA.	65
FIGURA 15 CUADRANTES GEOLOCALIZADOS QUE DIVIDEN LA LOCALIDAD DE JAÉN.	68
FIGURA 16 DIAGRAMA ENTIDAD RELACIÓN (DER).	69
FIGURA 17 DISTRIBUCIÓN DE LOS RECURSOS SANITARIOS URGENTES ACTIVADOS PARA ATENDER	

A PACIENTES CON ENFERMEDADES REPIRATORIAS, AGRUPADOS POR MESES DESDE 2012 HASTA 2019.96

FIGURA 18 DISTRIBUCIÓN DE LOS RECURSOS SANITARIOS URGENTES ACTIVADOS PARA ATENDER A PACIENTES CON ENFERMEDADES RESPIRATORIAS, AGRUPADOS POR DÍAS DE LA SEMANA DESDE 2012 HASTA 2019.98

FIGURA 19 COMPARACIÓN ENTRE EL VALOR REAL DE LAS DEMANDAS DE RECURSOS SANITARIOS URGENTES Y EL VALOR PREDICHO POR SVM CON KERNEL GAUSSIANO. VALORES POR SEMANA CONSIDERANDO SÓLO LOS DÍAS LABORABLES DE 2012..... 110

FIGURA 20 COMPARACIÓN ENTRE EL VALOR REAL DE LAS DEMANDAS DE RECURSOS SANITARIOS URGENTES Y EL VALOR PREDICHO POR SVM CON KERNEL GAUSSIANO. VALORES POR SEMANA CONSIDERANDO FINES DE SEMANA Y FESTIVOS EN 2012..... 114

FIGURA 21 EJEMPLO DE MODELO PREDICTIVO DE REGRESIÓN..... 120

FIGURA 22 PRINCIPALES VARIABLES CON PESO POSITIVO. 129

FIGURA 23 COEFICIENTES DE LA VARIABLE MES CON SVM KERNEL LINEAL. 131

INDICE DE TABLAS

TABLA 1 INFORMACIÓN PERSONAL DEL PACIENTE. ..60	60
TABLA 2 INFORMACIÓN SANITARIA DEL PACIENTE...61	61
TABLA 3 INFORMACIÓN CRONOLÓGICA DE LA ASISTENCIA.....63	63
TABLA 4 JAEN_V00_DEMANDAS.74	74
TABLA 5 JAEN_EX_INF_METEOROLOGICA.....75	75
TABLA 6 JAEN_EX_PROVINCIA.....76	76
TABLA 7 JAEN_EX_ESTACION_METEOROLOGICA.....76	76
TABLA 8 JAEN_V00_CUADRANTE_JAEN.78	78
TABLA 9 JAEN_V00_CALLES_CP.....78	78
TABLA 10 JAEN_V00_VALORES_X_CP.....82	82
TABLA 11 JAEN_V01_CALLES_DIRDT.82	82
TABLA 12 DÍAS CON INCIDENCIA EN EL AÑO 2012.87	87
TABLA 13 DÍAS CON INCIDENCIA EN EL AÑO 2015.90	90
TABLA 14 JAEN_V11_DEMANDA_FES_AG_CTRL.92	92
TABLA 15 JAEN_V12_DEM_AMP_FES_AG_CTRL.....93	93
TABLA 16 RECURSOS SANITARIOS DE URGENCIA ACTIVADOS PARA ATENDER A PACIENTES CON ENFERMEDADES RESPIRATORIAS, AGRUPADOS POR MESES DE 2012 A 2019.95	95
TABLA 17 RECURSOS SANITARIOS DE URGENCIA ACTIVADOS PARA LA ATENCIÓN DE PACIENTES CON ENFERMEDADES RESPIRATORIAS, AGRUPADOS POR DÍAS DE LA SEMANA DE 2012 A 2019.....97	97
TABLA 18 TABLA DE EJEMPLO DEL CONJUNTO DE DATOS UTILIZADO PARA GENERAR LOS MODELOS PREDICTIVOS.100	100

TABLA 19 MÉTRICAS DE ERROR DE LOS MODELOS PREDICTIVOS OBTENIDOS PARA LOS DÍAS LABORABLES UTILIZANDO 2012 COMO AÑO DE CONTROL..... 107

TABLA 20 PARÁMETROS DEL ALGORITMO SVM CON NÚCLEO GAUSSIANO UTILIZADO PARA GENERAR EL MODELO DE PREDICCIÓN. DATOS PARA LA PREDICCIÓN DE 2012..... 108

TABLA 21 ERROR ABSOLUTO MEDIO DE LOS RESULTADOS PARA LOS MODELOS DE DÍAS LABORABLES..... 110

TABLA 22 MÉTRICAS DE ERROR DE LOS MODELOS PREDICTIVOS OBTENIDOS PARA FINES DE SEMANA Y FESTIVOS UTILIZANDO 2012 COMO AÑO DE CONTROL. 112

TABLA 23 ERROR ABSOLUTO MEDIO DE LOS RESULTADOS DE LOS MODELOS DE FINES DE SEMANA Y FESTIVOS. 113

TABLA 24 COEFICIENTES UTILIZADOS PARA EL ALGORITMO SVM LINEAL 115

TABLA 25 PRINCIPALES VARIABLES CON PESO POSITIVO 129

TABLA 26 COEFICIENTES DE LA VARIABLE MES CON SVM KERNEL LINEAL 130

1. Introducción

La sanidad es uno de los sectores más sensibles de nuestra sociedad, y prueba de ello son los recursos y esfuerzos que se invierten en todo el mundo para intentar mejorar la gestión del sistema sanitario, especialmente en la optimización de los recursos sanitarios [1][2].

Aunque se ha producido un gran cambio en la forma de diagnosticar y tratar las enfermedades, apenas ha habido cambios en la forma de gestionar los servicios sanitarios en el siglo XXI.

En el campo del diseño de servicios han surgido numerosos estudios académicos, pero no abundan las investigaciones al respecto, sobre todo en el ámbito de los servicios sanitarios [3].

Una visión de conjunto del estado actual de la técnica en este ámbito muestra que la gran mayoría se orienta a lograr una mayor eficiencia económica en algunos aspectos del sector.

Por otro lado, existen trabajos científicos en los que se han ensayado diferentes técnicas para mejorar la gestión de los recursos sanitarios. Un claro ejemplo es el trabajo de Ramos y Cubillas, pionero en utilizar algoritmos predictivos para la gestión de la asistencia de los pacientes a los centros de salud [4]. Uno de los aspectos más importantes de este estudio es la demostración de la importancia de la inclusión de la componente espacial en los procesos predictivos. En este estudio podemos ver claramente el beneficio de la inclusión de la componente espacial en los resultados finales, con respecto a su no inclusión.

Además, el uso de métodos estadísticos para anticipar las tasas de llegada de pacientes a las organizaciones sanitarias permite programar el personal interno para satisfacer la demanda de servicios en función de la tasa de llegada de pacientes [5][6].

No obstante, esta investigación tenía la limitación de utilizar datos de unos pocos meses para extraer conclusiones sobre la llegada de pacientes. Este problema genera conclusiones menos fiables y más sujetas a la idiosincrasia a corto plazo de los datos, de ahí la importancia de trabajar con un buen histórico de datos, convenientemente revisados, estructurados y analizados antes de ser insertados en los modelos.

En definitiva, todos los tipos de investigación basados en modelos que proporcionan información anticipada sobre el comportamiento de un fenómeno como la demanda de recursos sanitarios son altamente dependientes de un gran volumen de datos disponibles con una alta calidad espacio-temporal.

Como introducción al entorno de nuestra investigación tenemos que indicar que en Andalucía el organismo encargado de gestionar y resolver las urgencias y emergencias sanitarias es el Centro de emergencias 061. La siguiente figura muestra la dotación de dicho centro en toda Andalucía.



Figura 1. Recursos Generales

Dichos recursos son gestionados desde un centro de coordinación en cada una de las provincias llamado Servicio provincial 061, como podemos ver en la Fig 2. Para dicha gestión en la actualidad se utiliza un sistema informático de gestión de emergencias en modo colaborativo llamado Centros en Red (CCR). Este sistema nos permite la gestión desde cualquier centro de Andalucía de todos o parte de los recursos de cualquier provincia.

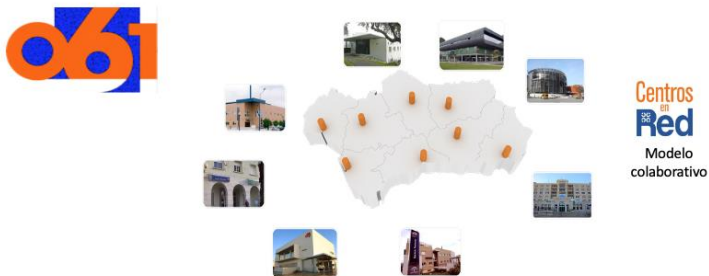


Figura 2. Servicios Provinciales CES 061.

El objetivo principal de nuestra investigación es el diseño y creación de una base de datos multidisciplinar, que contenga toda la información necesaria para poder aplicar sobre ella algoritmos de predicción que nos permitan conocer la necesidad de recursos que vamos a tener en una zona y fecha determinada.

Esta base de datos servirá como fuente para aplicar y analizar algoritmos de regresión de minería de datos con el fin de predecir la demanda de recursos de emergencia que se producirá en próximas fechas.

En la actualidad la dotación tanto de recursos humanos como materiales es fija, la ubicación de los mismos está predefinida y no está sujeta a variación. Pensamos que nuestra investigación puede ayudar enormemente en este aspecto, ya que, con antelación podremos conocer la necesidad con más exactitud de los recursos sanitarios que serán necesarios y su distribución por zonas. Con esta información se podría reorganizar el número de recursos y su ubicación de forma dinámica en función de las necesidades que nos indique nuestro sistema.

En la figura 3 podemos ver un esquema del proceso general.

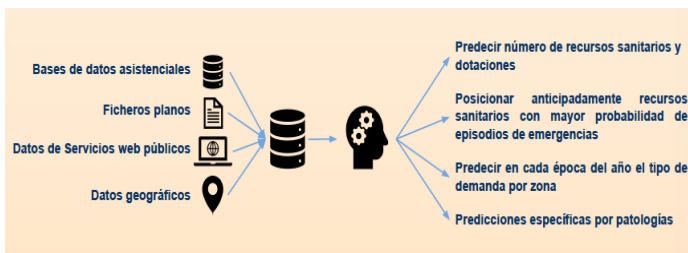


Figura 3 Proceso general.

En la anterior imagen podemos ver claramente los dos grandes bloques de nuestra investigación.

En el primero, el objetivo principal es conseguir una base de datos completa y optimizada. El nucleo principal lo compone toda la información sanitaria correspondiente a urgencias y emergencias de los últimos años recogida en el Centro de emergencias 061 de Jaén. Esta información será complementada con los datos meteorológicos de las zonas a estudiar, dicha información se ha obtenido en forma de ficheros planos que incorporaremos a nuestra base datos. Completando la información anterior incorporamos datos sociales y estadísticos obtenidos en las web de organismos públicos.

Toda esta información multiplica su valor añadiéndole a nuestra base de datos la componente espacial.

1.1 Estado del arte

No existen antecedentes relevantes de la implantación de sistemas que proporcionen una previsión diaria y suficientemente anticipada de la demanda de recursos en los servicios de emergencias sanitarias (es decir, que proporcionen una localización y gestión directa de los recursos que atienden diariamente a los pacientes) [7][8].

Existen algunos estudios que han puesto de manifiesto un importante incremento en la necesidad de optimizar la estructura de las bases de datos para hacer frente a la demanda de nuevos datos necesarios en la gestión sanitaria [9-12].

En la segunda década del siglo XXI se ha producido un aumento de las publicaciones destinadas a mejorar la estructura de las bases de datos adaptadas a la gestión diaria como forma de obtener información sanitaria más detallada. Proliferan nuevos canales de comunicación entre el paciente y los servicios sanitarios a través de dispositivos portátiles como tablets y smartphones, o mediante el uso del PC [13,14]

Posteriormente, la investigación en este sector ha mantenido esta línea de trabajo y aparecen nuevos retos en el horizonte. Además, la pandemia mundial iniciada en 2020 por el nuevo virus del síndrome respiratorio agudo severo (SARS)-CoV-2 (enfermedad por coronavirus 2019 [COVID-19]), ha provocado cambios drásticos en las prioridades sanitarias.

Las prioridades biomédicas han pasado a dominar la agenda, poniendo de relieve las lagunas de conocimiento multisectorial y los retos que hay que abordar para la gestión sanitaria de la pandemia.

Por el contrario, la gestión de la información y los sistemas de apoyo a la toma de decisiones pasaron a un segundo plano [15,16].

Sin embargo, una vez iniciado el proceso de inmunización de la población, es el momento de hacer balance y analizar cómo se han gestionado los recursos sanitarios y si, de alguna manera, una correcta toma de decisiones podría haber evitado la saturación de los servicios sanitarios. Son muchos y variados los datos que hay que valorar para llevar a cabo una adecuada gestión de los recursos sanitarios.

A pesar de la pandemia actual, la demanda de asistencia sanitaria sigue estando motivada por diferentes patologías y por diferentes razones.

En resumen, hoy en día se dispone de cada vez más datos, todos ellos procedentes de distintas fuentes, con diferentes formatos y diferente temporalidad y resolución. Por lo tanto, es necesario gestionar e integrar adecuadamente en las bases de datos centralizadas una gran variedad de datos procedentes de los nuevos dispositivos de entrada utilizados en la sanidad. También es importante que estas bases de datos sean capaces de integrar cualquier nueva variable resultante del progreso de la investigación sanitaria.

De este modo, la utilidad de la base de datos, además de suponer una herramienta eficaz de gestión de recursos y de aportar calidad al servicio, también tendría un papel importante en el seguimiento de enfermedades [17-20].

Este escenario requiere el desarrollo de herramientas y metodologías específicas orientadas a la consecución de estos objetivos de gestión sanitaria como Hamami et al. (2019) [21], quienes destacan que alcanzar el mejor modelo es una tarea compleja debido a la interacción de muchos componentes y a la variabilidad de los valores de los parámetros que conducen a dinámicas radicalmente diferentes. Por lo tanto, señala que el proceso de modelización puede mejorarse mediante el uso de técnicas de minería de datos [22][23].

Otro ejemplo del uso de técnicas de minería de datos en la gestión sanitaria para la toma de decisiones ya ha llegado a la conclusión de que pueden influir en los costes, los ingresos y la eficiencia operativa, manteniendo al mismo tiempo un alto nivel de atención al paciente [24]. Hay, por tanto, muchos aspectos que considerar y, sobre todo, hay que gestionar la gran

cantidad de datos que se generan cada día en torno a un servicio sanitario.

Por lo tanto, además de las herramientas de minería de datos, es un área importante de aplicación para los big data [25], que se conoce como big data médico [26].

Los macrodatos médicos proceden de diversas fuentes, como registros administrativos, historias clínicas, datos biométricos, datos de informes de pacientes, etc. También son de gran escala, se actualizan con gran rapidez, son polimórficos, incompletos y sensibles al tiempo [27]. Además, la cuestión de si los datos se utilizan adecuadamente o no sigue siendo un interrogante abierto.

El almacén de datos, más conocido como Data Warehouse (DW) es la respuesta al procesamiento de datos, pero las aplicaciones de los métodos tradicionales de DW en el ámbito sanitario requieren una atención considerable debido a la naturaleza empresarial única de este sector [28]. Guo, C., & Chen, J. (2023) [29] propusieron un enfoque basado en la utilización de Big data para el desarrollo de sistemas de información sanitaria.

El diseño de la base de datos proporciona la escalabilidad necesaria para cubrir otras aplicaciones especializadas sin necesidad de cambios estructurales a nivel de base de datos. La consecución de este objetivo requiere bases de datos con información administrativa y asistencial de varios años consecutivos [7,30,31,32], así como un modelo eficiente de almacenamiento y recuperación de big data sanitario para conseguir estimaciones válidas para una gestión óptima y de calidad [33,34].

En definitiva, para ofrecer una mejora en la calidad de la atención sanitaria, es imprescindible adaptar los sistemas de bases de datos para su uso en tecnologías DW y big data y en sus técnicas de explotación.

En el ámbito de las emergencias sanitarias, podemos citar el trabajo de Graham et al. 2018 [35] en el que se realizó un estudio predictivo del flujo de pacientes a urgencias desde los hospitales utilizando registros de dos grandes hospitales de la ciudad de Irlanda del Norte. Este trabajo alcanzó una fiabilidad superior al 80%. Otro ejemplo más reciente de predicción precoz del ingreso hospitalario de pacientes de urgencias nos lo da Kishore, K. et al. (2023) [36].

Otros estudios más recientes como los de Gurazada et al. (2022) [37] han realizado trabajos predictivos sobre la duración de la estancia de los pacientes en urgencias. A través de una revisión de la literatura se identificaron dieciséis factores potencialmente relevantes que impactan en los tiempos de espera.

Todos estos trabajos contribuyen a mejorar la atención al paciente al proporcionar a los gestores de los recursos sanitarios información anticipada. Sin embargo, en un servicio de urgencias se registra una gran cantidad de datos, no sólo de los pacientes, sino también del servicio prestado, los recursos utilizados y los factores externos en el momento de la urgencia.

Una vez más, queda patente que la correcta organización y almacenamiento de esta información heterogénea en una base de datos multiplica las posibilidades de extraer conocimiento oculto y capacidad predictiva de los datos. Presenta el diseño de una base de datos lo suficientemente compleja, es decir, con múltiples variables extraídas de cada demanda de emergencia sanitaria, para integrar todo tipo de información sanitaria que

permita un conocimiento avanzado y una mejor gestión de estos recursos, proporcionando así una atención de calidad al paciente.

1.2 Hipótesis y aportaciones de este trabajo

En la búsqueda y el estudio de los principales aspectos en la asistencia sanitaria de urgencias y emergencias, comprobamos que uno de los factores más importantes es el tiempo en el que el paciente recibe atención.

Los retrasos en la llegada al lugar donde se solicita la asistencia y en la recepción del tratamiento, pueden acarrear graves consecuencias negativas y un mal pronóstico para el paciente [38]. La vida del paciente y sus posibles secuelas dependen directamente de este intervalo de tiempo.

Diversos estudios médicos destacan la importancia de este factor y enumeran las diversas complicaciones que aparecen en los pacientes a medida que avanza el tiempo de espera para recibir asistencia [39, 40].

Los motivos de la demora en el tiempo de atención al paciente son diversos, desde los inherentes a la población que solicita atención, como las características demográficas, el nivel socioeconómico, la etnia, etc. [42, 43], hasta el tipo de patología. Sin embargo, los factores relacionados con los sistemas sanitarios también desempeñan un papel importante. Entre ellos se incluyen la disponibilidad de servicios, la accesibilidad de los centros sanitarios, la aceptabilidad y la adecuación de los recursos hospitalarios [44]. La interacción de factores individuales y externos puede provocar más retrasos en la llegada y el tratamiento.

Uno de los problemas más importantes que suelen producirse en una emergencia es la saturación de recursos en la zona donde se produce el suceso, lo que provoca un retraso en la atención. Por ello, un factor crucial para solucionar este problema es disponer de recursos sanitarios correctamente dimensionados para minimizar el riesgo que conlleva este tipo de situaciones.

Los responsables de gestionar la disponibilidad de estos recursos deben tener en cuenta múltiples factores que hay que considerar para ofrecer un servicio sanitario de urgencia adecuado. Entre ellos, la experiencia de gestión del responsable del servicio, que normalmente se basa en datos de series temporales de años anteriores, como puede ser la media de demandas sanitarias en la misma época de años anteriores. En este sentido, hay que tener en cuenta que los motivos de las demandas sanitarias son múltiples, por lo que se produce una fluctuación en el número de demandas que se producen. Por este motivo, para gestionar adecuadamente los recursos, también es importante tener una previsión del tipo de demandas sanitarias que se pueden producir.

Para obtener esta información predictiva, es muy útil utilizar los recientes avances en Artificial Intelligence (AI) y Machine Learning (ML) (Aprendizaje Automático). Esta tecnología ha permitido avances sustanciales en la predicción e identificación de emergencias sanitarias, poblaciones enfermas, estado de enfermedad y respuesta inmune, entre otros [41].

Estos datos se utilizan para generar el modelo predictivo. En el campo de la medicina existen numerosas investigaciones en las que se han utilizado estos algoritmos para predecir aspectos

clave de la salud, como la gestión de citas, predecir el número de pacientes que acuden a los centros de salud, etc. [42, 43].

Existen diferentes algoritmos aplicables a este tipo de tareas, cada uno de los cuales es más efectivo en función del tipo de datos a predecir. De hecho, hay autores que han realizado trabajos sobre la predicción de enfermedades infecciosas y la atención sanitaria de urgencia que se han derivado por este motivo. En estos estudios se han utilizado redes bayesianas para obtener los datos de predicción [44].

Otros estudios han utilizado algoritmos de clasificación en los servicios de urgencias hospitalarios para realizar una clasificación inicial de los pacientes en función de su situación médica. Incluso en el caso del suministro de medicamentos en centros sanitarios, es importante disponer de información anticipada sobre qué pacientes necesitarán qué medicamentos. En este campo, también hay estudios que han inferido datos sobre la cantidad de demandas de fármacos relacionados con enfermedades de insuficiencia respiratoria [45].

La creciente disponibilidad de datos clínicos, de movilización de recursos y otros datos externos de múltiples fuentes hace que el uso de técnicas computacionales como el ML permita el procesamiento significativo de grandes cantidades de datos complejos, contribuyendo así a la generación de sistemas expertos capaces de anticipar situaciones críticas en urgencias.

Una vez consigamos nuestro principal objetivo de crear la base de datos multifuente, un segundo objetivo será su aplicación en el servicio de urgencias y emergencias de la ciudad de Jaén, en el sur de España. Este segundo objetivo será una aportación fundamental en la gestión real de las urgencias y emergencias.

Teniendo en cuenta que la zona geográfica en la que estamos, es la zona con mayor densidad de olivos del mundo, podemos prever una gran acumulación de problemas respiratorios en la población debido al polen del olivo, entre otras especies.

Por otro lado, después de analizar las diferentes patologías atendidas en el Servicio de Urgencias, se ha observado que un número significativo, el 10 %, están causadas por enfermedades respiratorias y se sabe que éstas, además de por el factor anteriormente citado, se ven influenciadas por diferentes factores: contaminación, temperatura, humedad, etc. [46-48].

Por los motivos explicados anteriormente, centramos la primera aplicación de nuestra investigación en la gestión de los recursos de los pacientes afectados por patologías respiratorias. Nuestro sistema obtendrá una predicción del número de demandas que se producirán en un futuro inmediato. Se genera así una herramienta útil para los responsables de dimensionar estos servicios sanitarios, aportando datos que ayuden a la toma de decisiones en esta tarea.

De hecho, hay varias investigaciones en las que se ha utilizado el ML en el ámbito sanitario para ayudar a un médico o analista a realizar sus funciones, identificar tendencias sanitarias y desarrollar modelos de predicción de enfermedades [49-53]. Otros han utilizado estas técnicas para gestionar agendas de citas en centros de salud de atención primaria [42,43,54].

La aportación principal de nuestra investigación radica en que reduciremos el tiempo de espera para la asistencia a los pacientes, ya que al conocer la predicción de demanda de recursos sanitarios para urgencias y emergencias, podríamos reubicar los recursos sanitarios disponibles en las zonas con

más incidencias. Así como ampliar los recursos en las zonas y fechas en las que nuestro sistema prediga una mayor actividad.

Hay que tener en cuenta que las patologías que se gestionan en urgencias y emergencias son tiempo-dependientes. Cada periodo de tiempo recortado en la respuesta a estas peticiones deriva directamente en un beneficio directo en la salud del paciente.

Actualmente en el núcleo urbano de Jaén existe una UVI móvil 061 y diversos recursos de distinta índole que actúan en el caso de que dicha UVI esté ocupada y sea necesario un recurso de este tipo. Por este motivo es fundamental que la ubicación de esta UVI y los distintos recursos sea la zona de mayor incidencia en cada periodo del año. Así como un incremento de dichos recursos en los periodos en los que nuestro sistema prediga más actividad de este tipo. Por ejemplo épocas de alergia o de aumento o disminución de la población.

2. Metodología

En este apartado explicaremos los diferentes aspectos que rodean a nuestra investigación. Analizaremos en qué contexto geográfico vamos a ubicar nuestro estudio para justificar las distintas vías de investigación tomadas. Posteriormente analizaremos la metodología de trabajo utilizada en urgencias y emergencias en dicho entorno geográfico.

Analizaremos la metodología utilizada para la integración y gestión de toda la información en nuestro sistema para poder dar respuesta a nuestro objetivo.

2.1 Área geográfica de la investigación

El municipio de Jaén está situado en la provincia de Jaén, en la comunidad autónoma de Andalucía, España. Está situado en el sur del país, figura 4. Jaén está situado en una colina, lo que le confiere un relieve montañoso. Se encuentra en la confluencia de los ríos Guadalbullón y Guadalquivir.

El entorno natural que rodea Jaén es característico de la zona mediterránea, con predominio de monte bajo y olivares. Los olivares forman parte integrante del paisaje de la región, y la producción de aceite de oliva es una actividad económica importante en la zona. Precisamente la floración de este cultivo provoca numerosos episodios de asma entre la población.

La provincia de Jaén cuenta con más de 70 millones de olivos, lo que, durante la época de floración, da lugar a

concentraciones de polen muy elevadas, que causan grandes daños a la salud de la población alérgica.

En cuanto a las condiciones meteorológicas en Jaén, la zona experimenta un clima mediterráneo. Los veranos suelen ser calurosos y secos, con temperaturas que a menudo superan los 30 grados Celsius (86 grados Fahrenheit). Los inviernos son suaves, con temperaturas que oscilan entre los 8 y los 15 grados Celsius (46 y 59 grados Fahrenheit). Las precipitaciones son relativamente escasas a lo largo del año, con lluvias ocasionales sobre todo en otoño y primavera.

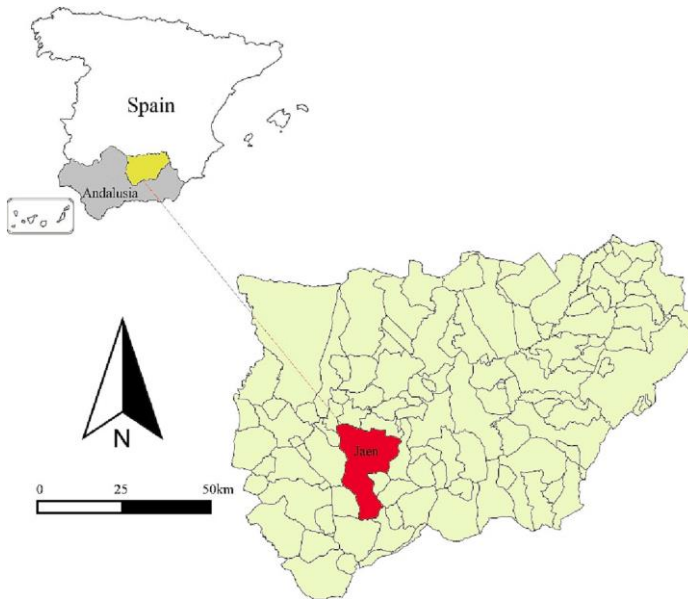


Figura 4 Localización de la zona de estudio en Jaén, Andalucía (España).

En la ciudad de Jaén el Centro de emergencias sanitarias 061 es el organismo que tiene las competencias en la coordinación y gestión de urgencias y emergencias [55].

El Centro de emergencias sanitarias tiene funcionando de forma permanente, 2 Equipos de Emergencias Terrestres (EET) y 1 Equipo de Coordinación Avanzada (ECA). Un EET está ubicado en el núcleo urbano de Jaén, un segundo EET tiene su base en la localidad de Bailén y el ECA en Úbeda.

Cada equipo tiene asignada una zona de cobertura que comprende la localidad en la que se ubica y las localidades limítrofes. Todas las zonas que se quedan fuera de estas coberturas son atendidas a nivel de emergencias por 2 Equipos de Emergencias Aéreas (EEA). Un helicóptero ubicado en la ciudad de Córdoba y que cubre la mitad occidental de la provincia de Jaén y un segundo helicóptero ubicado en la localidad de Baza y que da servicio a la mitad oriental.

Cada recurso EET está dotado tres categorías de profesionales:

- 1 Técnico de Emergencias Sanitarias (TES)
- 1 Diplomado Universitario en Enfermería (DUE)
- 1 Médico

Cada recurso ECA está dotado de:

- 1 Técnico de Emergencias Sanitarias (TES)
- 1 Diplomado Universitario en Enfermería (DUE)

Tal y como muestra la figura 5, en la provincia de Jaén el Centro de emergencias sanitarias 061 tiene un total de 57 profesionales sanitarios:

- 15 TES
- 25 DUE
- 17 Médicos



Figura 5. Plantilla asistencial por provincias.

Tal y como hemos indicado anteriormente en la ciudad de Jaén existe un EET que cubre la ciudad de Jaén y las localidades más cercanas. En el caso de nuestra investigación nos centraremos en el núcleo urbano de Jaén.

Un aspecto importante a resaltar, y que en el futuro veremos que es importante en nuestra investigación, es el tiempo de respuesta de este equipo de emergencias en sus asistencias en el núcleo urbano de Jaén. Tal y como se muestra en la Fig. 6 en el año 2022 el tiempo medio de respuesta de este equipo de emergencias en sus asistencias en la ciudad de Jaén fue de 9 minutos y 31 segundos.

Como apunte adicional indicar que en cada uno de esos casos el centro de coordinación invirtió una media de 4:40 minutos

en coordinar dicha demanda, es decir, atender la llamada, decidir el recurso más idóneo y dar los consejos pertinentes.

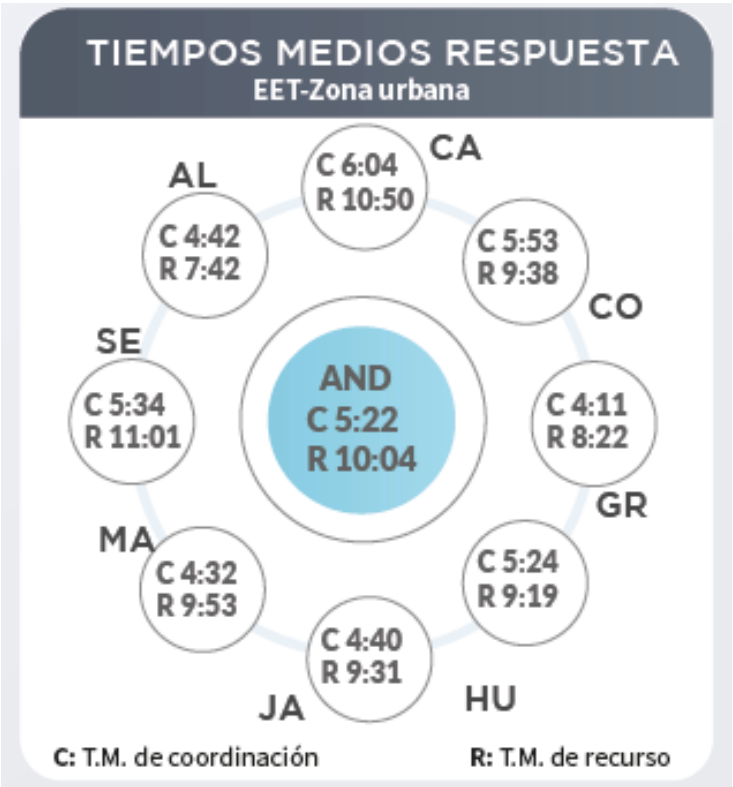


Figura 6. Tiempos medios de respuesta en zonas urbanas en Andalucía.

En el año 2022 los equipos sanitarios del 061 de Andalucía atendieron a 61.192 pacientes. La Fig. 7 nos muestra la

distribución de todos estos pacientes atendidos por tipo de equipo de emergencias.



Figura 7. Pacientes atendidos por equipos 061 en Andalucía.

2.2 Descripción de la metodología de Trabajo en el área de Emergencias en Andalucía.

Se define emergencia como aquella situación en la que la vida de una persona corre peligro, en caso contrario, se identifica como urgencia. Actualmente, existe un número de teléfono de emergencias a disposición de los ciudadanos en Andalucía (061) y un número de urgencias (902505061). Como se ha explicado anteriormente, este servicio es prestado por el CES (Centro de Emergencias Sanitarias 061), que cuenta con ocho servicios provinciales en Andalucía, uno por provincia.

Los servicios provinciales son las sedes desde las que se gestionan todas las urgencias y emergencias de cada provincia. El núcleo más importante de cada servicio provincial es la sala de coordinación, donde se reciben todas las llamadas realizadas por los ciudadanos al número de emergencias de cada provincia.

Cada sala de coordinación está formada por uno o varios médicos coordinadores y por un conjunto de gestores telefónicos que gestionan la demanda sanitaria del ciudadano. Estos gestores telefónicos atienden al ciudadano recabando toda la información posible sobre la demanda en curso, el médico coordinador participa en esta gestión y finalmente decide, en función de la gravedad del paciente, qué recurso se moviliza para resolver la demanda.

Desde la sala de coordinación se movilizan todos los recursos sanitarios necesarios para todas las urgencias y emergencias de la provincia. CES cuenta con recursos propios como los equipos de emergencias terrestres (UVI móviles) y los equipos de emergencias aéreas (Helicópteros Sanitarios), además de coordinar y movilizar todos los recursos de emergencias del Servicio Andaluz de Salud (SAS) y todas las ambulancias de la red de transporte urgente (RTU) de la provincia.

Existen aproximadamente 60 unidades en el Servicio Provincial de Jaén. Una vez activados los recursos desde la sala de coordinación, se dirigen al lugar de la asistencia, quedando registrados en el sistema informático todos los movimientos de dichas unidades, conociendo en tiempo real la geolocalización de todas ellas y el momento exacto de resolución de la asistencia donde se encuentra.

Todas las unidades de urgencias y emergencias disponen de dispositivos electrónicos (tablets) en los que se registra la historia clínica del paciente de la asistencia prestada (HCDM Historia Clínica Digital en Movilidad), que se informatiza al mismo tiempo que la resolución. La información más relevante de esta historia es enviada al centro coordinador para su almacenamiento, junto con la demanda creada por el gestor

telefónico al inicio del proceso, finalizando esta demanda y dando inicio al ciclo, figura 8.

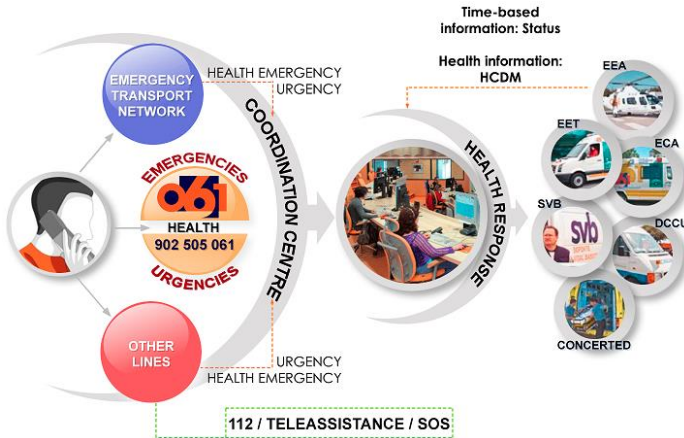


Figura 8 Ciclo completo de gestión de urgencias y emergencias.

Las unidades anteriormente indicadas se encuentran en ubicaciones preestablecidas y estáticas. Estas unidades tienen una base de emergencias desde la que se desplazan a la asistencia, esta base es siempre la misma para cada equipo y no cambia en ningún momento del año, no adaptándose a la realidad sanitaria de cada momento. Una mejora importante en este sistema sería cambiar la posición de los recursos disponibles de forma dinámica, según el tipo de asistencia que prestan, el número de ellos y la época del año en que están disponibles. Para ello, la base de datos presentada en nuestra tesis puede suponer un importante avance, ya que el diseño, la estructura y el nivel de detalle de la información almacenada permiten alcanzar este objetivo. Conseguiríamos dinamizar y

adaptar un servicio sanitario a la necesidad real del usuario en cada momento y situación, aumentando la calidad asistencial.

2.3 Objetivos de la tesis

Principales objetivos de la tesis:

- Predicción del número de recursos sanitarios y dotaciones:

Como indicamos con anterioridad los recursos sanitarios y la dotaciones que los componen están predefinidos y no cambian su constitución ni ubicación. Nuestro trabajo permitiría una actualización permanente de dichos recursos y dotaciones a la forma más eficaz y necesaria en cada momento, dependiendo de los factores externos que componen la realidad.

- Posicionamiento dinámico de los recursos.

Como indicamos con anterioridad, en la actualidad, la ubicación de los recursos en cada area sanitaria es fija, nuestra investigación permitiría ir actualizando esa ubicación teniendo en cuenta los aspectos externos influyentes en cada momento, como por ejemplo, la meteorología, el incremento de población, etc...

- Predicción del tipo de demanda por zona en cada época del año.

El sistema de predicción generado por nuestra investigación nos indicará en cada época del año que tipo de patología será la más demandada en cada zona. Para ello tendremos en cuenta además de los datos históricos sanitarios, parámetros como temperaturas, humedad, orografía, etc... Esto nos permitirá dotar a los equipos de emergencias del material necesario para hacer frente a estos tipos de demandas.

- Predicción específica por patologías.

Teniendo en cuenta las patologías que dan lugar a un mayor número de demandas, nuestra investigación puede ser utilizada por profesionales sanitarios para diseñar nuevos modelos específicos de funcionamiento para conseguir una asistencia más adecuada a dichos pacientes.

2.4 Integración y gestión de datos

El conjunto de datos necesario para esta investigación incluye toda la información correspondiente a las demandas asistenciales recibidas por los servicios de emergencias de la ciudad de Jaén, en el sur de España, desde 2012 hasta 2019 inclusive.

En este sentido, es importante tener en cuenta que los datos están geolocalizados, es decir, se ha tenido en cuenta su componente espacial, es decir, dónde se ha producido la demanda de asistencia sanitaria urgente.

Además, se han incluido una serie de datos secundarios necesarios para contextualizar las condiciones particulares en las que se produce cada episodio de emergencia. Así, se han

añadido diversas variables multifuente, como datos medioambientales y meteorológicos de la zona en los años anteriormente indicados, así como datos socioeconómicos de cada una de las zonas en las que se divide la ciudad de Jaén [56].

A continuación se describen cada una de las variables que componen el conjunto de datos:

- Datos de urgencias sanitarias: Se almacena toda la información relacionada con la asistencia sanitaria que ha recibido el paciente, desde que se recibe la llamada telefónica en el centro coordinador, hasta que el equipo médico da por finalizado el caso. Estos datos incluyen las peticiones de asistencia del usuario en situaciones de urgencia y emergencia, incluyendo diagnóstico, prueba clínica, tratamiento, antecedentes, recursos movilizados y tiempos de actuación detallados, así como la geolocalización y su resolución.
- Datos atmosféricos: Datos recogidos por la Red de Información Ambiental de Andalucía (REDIAM) y facilitados por la Consejería de Medio Ambiente y Ordenación del Territorio [57]. Estos datos incluyen valores de temperatura, precipitación, humedad, velocidad del viento y radiación solar.
- Datos sociológicos: Los datos relacionados con la información personal de los usuarios se dividieron en:
 - Nivel económico de los pacientes de cada zona: Analizando, por un lado, el valor catastral de los inmuebles, extraído de la página web de la Dirección General del Catastro de España [58]. También se añadió un análisis del precio actual de la vivienda en cada distrito de la ciudad a través de portales web inmobiliarios.

- Nivel de desempleo de los pacientes y sus unidades familiares: Esta variable se obtuvo a partir de los datos proporcionados por el Instituto Nacional de Estadística [59]. Esta institución pública nos proporcionó el tipo de población de cada sección censal.

- Nivel de estudios, edad de los ciudadanos y miembros de la unidad familiar: Estos datos se obtuvieron de la página web del Instituto de Estadística y Cartografía de Andalucía [60]. La metodología seguida en este trabajo se secuencía en varias fases, desde la comprensión de los datos, su preparación y la generación del modelo hasta su validación. En la figura 9 podemos ver las primeras fases de nuestra investigación.



Figura 9 Construcción Base de Datos relacional multimodal.

Como podemos ver en la figura 10, la segunda parte de nuestra investigación consiste en la aplicación de sistemas predictivos capaz de ofrecernos resultados aplicables en el mundo real.

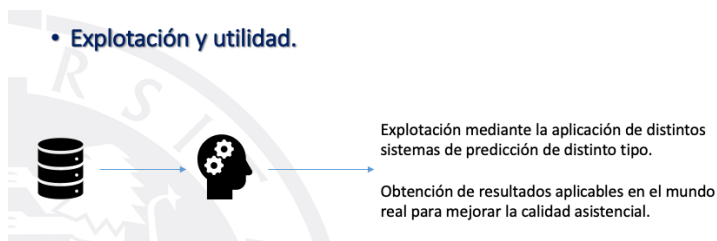


Figura 10 Explotación y utilidad.

En nuestro caso particular, el atributo objetivo de nuestro modelo, es el número de demandas sanitarias urgentes por enfermedades respiratorias que se recibirán en el centro de control para una fecha seleccionada.

El objetivo es predecir un dato desconocido, pero que se puede deducir a partir de otras variables conocidas y que están directa o inversamente relacionadas con el atributo objetivo.

En este sentido, el estudio se basa en un análisis de minería de datos supervisado, en el que el valor de una variable desconocida se deduce a partir de unas pocas variables conocidas.

El flujo general de la metodología llevada a cabo consta de algoritmos y técnicas de ML utilizados de la siguiente manera en cada fase:

1. Extracción y carga de datos. Los datos meteorológicos se descargan de servidores web públicos. Los datos relativos a la asistencia sanitaria urgente prestada son suministrados por

Centro de Emergencias Sanitarias. Ambos se cargarán en el sistema de gestión de bases de datos.

2. Primer análisis de los datos. Se exploran y analizan todos los datos utilizando técnicas de distribución (histogramas), con el objetivo de revisar los datos y depurar aquellos cuya dispersión o variabilidad pueda causar inconsistencias en el estudio.

3. Detección de anomalías. La detección de anomalías se implementa como un algoritmo de clasificación de clases capaz de predecir, con cierta probabilidad, si un registro de los datos es típico de la distribución. El objetivo de esta fase es identificar aquellos casos que no son comunes en nuestros datos.

4. Transformaciones de datos. En esta sección se transforman tanto los datos de rendimiento como los meteorológicos, es decir, se adaptan los formatos, las unidades, se reescalan, etc., para que puedan ser explotados de forma óptima por los modelos predictivos.

5. Agregación de datos. La información descargada se agrega mensualmente, por lo que el resto de datos que se añadan al estudio deberán agregarse de la misma forma.

6. Integración de datos. Para nuestro trabajo contamos con información heterogénea procedente de distintas fuentes. Así, para llevar a cabo el estudio de minería de datos, es necesario integrar toda la información en una única fuente que sirva de entrada para los modelos predictivos.

7. Detección del nivel de influencia de los atributos de entrada sobre el atributo objetivo. Antes de generar el modelo, se analiza la influencia de cada atributo sobre el atributo objetivo (número de demandas sanitarias urgentes) para incluir o

excluir atributos del estudio en función de su nivel de influencia en la predicción.

8. Aplicación de algoritmos de regresión [61-63]. Se prueban distintos algoritmos de regresión para predecir la asistencia sanitaria. El objetivo del análisis de regresión es determinar los valores de los parámetros de una función que mejor se ajuste a un conjunto de datos observacionales. Existen diferentes familias de funciones de regresión y diferentes formas de medir el error. En este trabajo hemos analizado funciones de regresión lineales y no lineales, así como distintos parámetros para evaluar la bondad de ajuste de los modelos.

2.5 Algoritmos de Minería de Datos

Como se indicó en la introducción, en este trabajo se utilizaron técnicas de minería de datos, basadas en los atributos almacenados en la base de datos diseñada. Previo al desarrollo de los modelos, se revisaron las siguientes técnicas:

El estudio predictivo realizado en este trabajo se basó en el uso de algoritmos de regresión. Entre ellos, la regresión lineal, la regresión logística, el modelo de regresión generalizada, la máquina de vectores soporte (SVM) de una clase, etc.

En este estudio, se desconocía a priori la naturaleza de las variables, que, como ya se ha dicho, son de naturaleza heterogénea. Se utilizaron algoritmos de regresión lineal y no lineal.

2.5.1 Algoritmos de regresión lineal

Los algoritmos puramente lineales tienen una gran fuerza debido a sus características y simplicidad ya que se calculan con una simple suma ponderada de las variables:

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p + \epsilon$$

El primer algoritmo seleccionado en este estudio fue el de los modelos lineales generalizados (MLG), que funciona matemáticamente como la suma ponderada de las características con el valor medio de la distribución asumida por la función de enlace g , que puede elegirse de forma flexible en función del tipo de resultado.

$$g(EY(y|x)) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

$E(Y)$ es el valor esperado de Y .
 g es la función de enlace.

En otras palabras, este algoritmo es una extensión de los algoritmos lineales que permite modelizar distribuciones lineales o normales y varianzas no constantes.

Los modelos lineales hacen una serie de suposiciones restrictivas, en las que el objetivo se distribuye normalmente condicionado al valor de los predictores con una varianza constante, independientemente del valor de la respuesta predicha. En este sentido, GLM relaja estas restricciones, y para un ejemplo de respuesta binaria, la respuesta es una probabilidad en el rango $[0, 1]$ [64, 65].

Otro algoritmo lineal seleccionado fue SVM, que tiene la ventaja de poder utilizarse con diferentes kernels. Los kernels permiten distribuir los datos en un hiperplano según una función, lo que facilita la adaptación del algoritmo a la naturaleza de los datos, permitiendo infinitas transformaciones.

En este estudio se ha trabajado con la SVM con kernel lineal, figura 11. Cuando se utiliza el kernel lineal, se realiza la siguiente transformación

$$K(x, x') = x \cdot x'$$

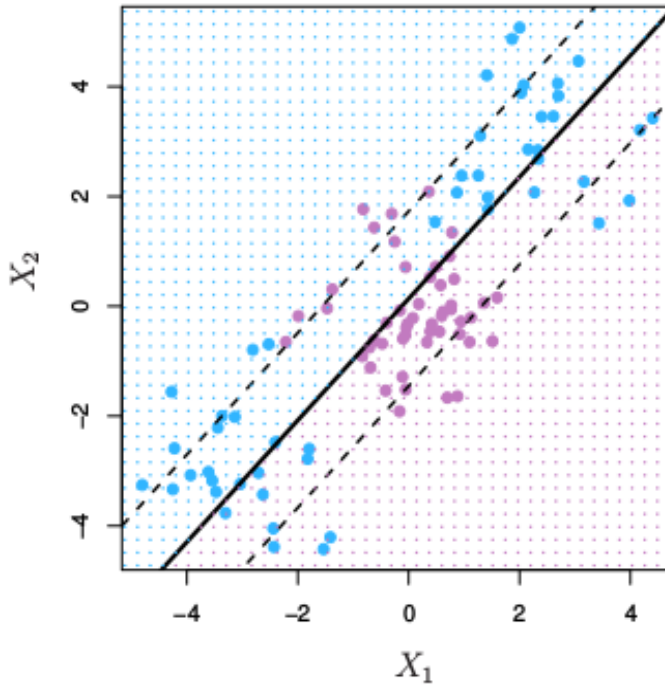


Figura 11 SVM Kernel Lineal.

Este algoritmo tiene la ventaja de que se adapta muy bien si la naturaleza de los datos es lineal, y si hay muchas variables predictoras (como en el caso estudiado).

Hay que tener en cuenta que en este algoritmo, no hay límite superior en el número de atributos predictores, y las únicas limitaciones son las impuestas por el hardware.

2.5.2 Algoritmos de regresión No Lineal

En este caso, el algoritmo SVM se aplica con un kernel gaussiano.

Este kernel aplica la siguiente transformación a los datos:

$$K(x, x') = \exp(-\gamma \|x - x'\|^2)$$

Podemos ver un ejemplo en la figura 12.

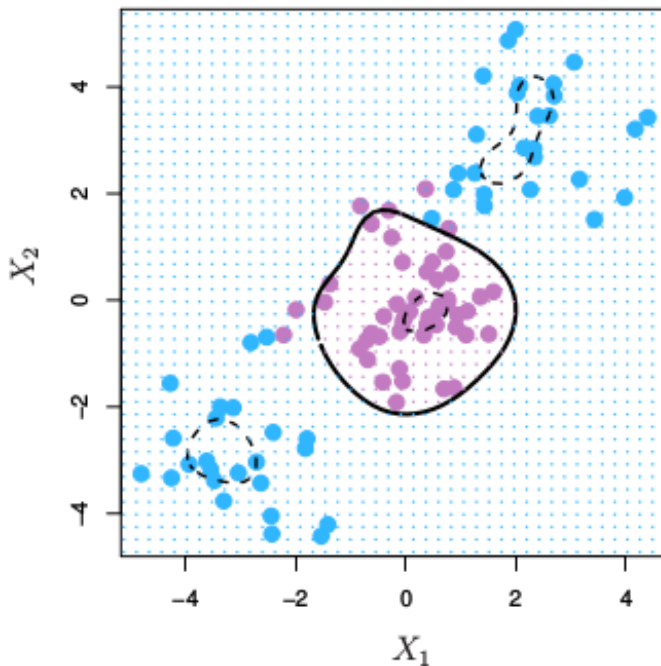


Figura 12 SVM Kernel Radial.

El valor de γ controla el comportamiento del kernel. Cuando es muy pequeño, el modelo final es equivalente al obtenido con un kernel lineal, a medida que aumenta su valor, los datos se alejan, formando una campana de Gauss en el hiperplano, ajustándose muy bien cuando la naturaleza de los datos no tiene una distribución lineal.

La separación lineal se entiende fácilmente cuando se estudia en espacios de 2 o 3 dimensiones.

En 2 dimensiones, que dos grupos de observaciones sean linealmente separables significa que existe al menos una recta que permite separar perfectamente los dos grupos.

En el caso de 3 dimensiones, los datos son linealmente separables si existe un plano tal, que es capaz de separar perfectamente los grupos. Este mismo concepto puede generalizarse para cualquier número de dimensiones, la condición sigue siendo la misma, que exista un elemento de una dimensión menor que separe los grupos. Ha este elemento se le conoce como hiperplano.

Utilizando el Kernel Gaussiano RBF podemos pasar los puntos de un plano 2D a un plano 3D simplemente desplazando todos los puntos verdes por encima de los rojos usando una función de mapeo como la RBF gaussiana, que introduce un hiperplano en el espacio dado, como se muestra en la figura 13.

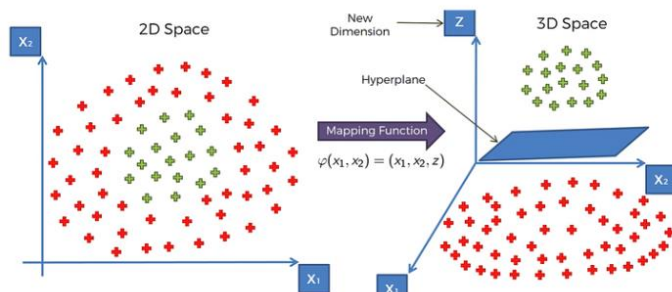


Figura 13 SVM Kernel Gaussiano RBF, paso de 2D a 3D

En resumen, estas son las ventajas de estos tres algoritmos en este estudio, partiendo de la hipótesis de desconocimiento a priori de la relación entre las variables con el atributo objetivo y considerando también que los datos de entrenamiento con los que contábamos eran limitados y las variables predictoras numerosas.

Además, la complejidad de estos algoritmos hace que la relación entre los atributos utilizados no pueda describirse mediante una ecuación específica.

En nuestro estudio se aplicaron los siguientes algoritmos:

- Algoritmo de longitud mínima de descripción (MDL) [66]: Para la detección de la importancia de los atributos, se descartan todos los atributos que no guardan relación con el atributo objetivo.
- Algoritmos de regresión: Una vez conocidos los atributos válidos, se prueban varios algoritmos para determinar cuál de ellos tiene mejor precisión. Se utilizan algoritmos de modelos lineales generalizados (GLM) [64] y máquinas de vectores de soporte (SVM) (kernel lineal y gaussiano) [67]. El algoritmo

GLM es un modelo lineal puro. Por su parte, las máquinas de vectores soporte (SVM) son un potente algoritmo basado en la teoría del aprendizaje estadístico.

La principal ventaja del algoritmo SVM es que se puede configurar con diferentes kernels, en este caso, hemos utilizado un kernel lineal y gaussiano.

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

3. Diseño de la investigación

La investigación tienes dos grandes partes, la primera es todo lo concerniente a la base de datos. Detallaremos todos los pasos seguidos en su diseño, creación, depuración y optimización para su objetivo final.

El segundo bloque que analizaremos será el diseño, creación y ejecución de un modelo predictivo que utilice la base de datos anterior. En nuestro caso nos centraremos en la optimización de recursos sanitarios en casos de dificultad respiratoria.

3.1 La base de datos

El primer objetivo intermedio que nos planteamos en nuestra investigación es la construcción de la base de datos. Dicha base de datos será la base sobre la que aplicaremos los distintos modelos predictivos en la búsqueda de nuestro objetivo final.

El flujo de trabajo del diseño de la base de datos se dividió en tres fases.

- 1.- Recopilación de datos de todas las fuentes utilizadas en nuestro sistema.
- 2.- Depuración de los datos.
- 3.- Estructuración y preparación de la base de datos para su explotación.

Proceso de limpieza de datos para facilitar su gestión y análisis. En esta fase, se llevó a cabo un proceso de limpieza, normalización y agrupación. Se partió de la tabla original, demandas, que tenía 84 atributos, en la que se encuentran todos

los detalles de las demandas de asistencia realizadas por los equipos de emergencia.

Con el fin de preparar la estructura de la base de datos para diferentes tipos de explotación, esta información se reestructuró en bloques específicos que incluyen varias tablas.

3.1.1 Bloques de información integrados

3.1.1.1 Bloque de información sanitaria

Este es el bloque más importante de la base de datos ya que son tablas que contienen los datos sanitarios. Estas tablas contienen toda la información correspondiente a la asistencia prestada por los equipos de emergencias durante los últimos años en la provincia de Jaén.

3.1.1.1.1 Datos personales del paciente

En este bloque se incluyen los campos de información del paciente que contienen la información personal de cada paciente. Los campos más significativos son la edad, el sexo, la fecha de nacimiento y la dirección, tabla 1.

Attribute	Description
Age	Patient 's years old
Sex	Man/Woman
Birth	Date of birth
Id Province	ZIP code
Id_District	District code
Id_Town	Town code
Id_Street	Street code
Id_Number	Number
Id_Door	Door

Tabla 1 Información personal del paciente.

3.1.1.1.2 Información sanitaria del paciente.

Contiene toda la información correspondiente a la salud del paciente atendido. Los más significativos se muestran en la siguiente tabla 2.

Attribute	Description
ldTypeofdemand	Classification of demand Type of demand: A Attendance / T Transport / I Informative
ldTypeofdemand1	First level of detail of the type of demand. E.g: 01 Transport accident 02 Alteration of vital signs 04 Dyspnoea ...
ldTypeofdemand2	Second level of detail of the type of demand
ldTypeofdemand3	Third level of detail of the type of demand
ldSResource	Resource Code
ldAssistance	Attendance Code
ZipCode	Zipcode
ClinicalJudgment 1	First Clinical Trial
ClinicalJudgment 2	Second Clinical Trial
ClinicalJudgment 3	Third Clinical Trial
ldResolveCode	Resolution Code: 1* do not arrive to see patient 2* Arrive but do not act 3* Attend to patient
ldFinance	Financing code. (State funding, Private companies...)
AdrDTDestination	Destination code of the resource.
Situation	Team attending: U Urban P Peripheral, if the assistance is covered by either of the two teams.
AdrDTSituation	Equipment Coverage Zone: Urban I Peripheral
IdAdmissionCenter	Hospital Admission Centre

Tabla 2 Información sanitaria del paciente.

3.1.1.1.3 Información cronológica de la asistencia

En relación con la información sobre la ambulancia movilizada para cada asistencia, se registra la hora de inicio y fin de cada intervalo de asistencia.

Esta información incluye la hora a la que el recurso móvil es activado por el centro coordinador, la hora a la que llega al lugar de la asistencia médica, el tiempo transcurrido durante la actuación sobre el paciente, la hora de transporte del paciente al hospital y la hora a la que el recurso móvil vuelve a estar disponible para la siguiente asistencia. Tabla 3.

Attribute	Description
Year D	Year of date of attendance.
Month D	Month of date of attendance.
Day D	Day of date of attendance.
IdProvince	Province code.
Requestdate	Request date
IdRequest	Request code
IdCall	Call code
IdLine	Line code
IdLineType	Line type code
IdAlertant	Alert source code (User, General emergency service...)
inLetTime	Incoming call time
RingTime	Time a1 which the system rings
AnswerTime	Time the call is answered
ResourceCreationTime	Time of resource creation
ActivationTime	Resource Activation Time
ExitTime	Time of departure of the resource
AtrivalTime	Time of arrival of the resource
Load Time	Time of patient loading in the ambulance (resource)
DestinationTime	Time of arrival at destination
OperationTime	Time at which the resource becomes operational
AvailableTime	Time when you are fully available for other assistance
CoordinationTime	Demand coordination time
ActivationTiming	Time taken to activate the resource
AttentionTiming	Patient care time
AnswerTiming	Response time (from the time the call comes in until the resource reaches the patient)
IdExclusionGround	Reason for exclusion in case of failure to send the appropriate resource for that request
IdResourceUnit	Resource Unit sent.
IdResourceType	Resource Unit sent.

Tabla 3 Información cronológica de la asistencia.

3.1.1.2 Información meteorológica

Numerosos estudios indican que los factores ambientales y meteorológicos influyen directamente en afecciones como las alergias e influyen directamente en la aparición de determinadas enfermedades como las alergias o determinadas enfermedades crónicas.

Tenemos que tener en cuenta que Jaén es la mayor productora de aceite del mundo y, por tanto, la floración del olivo en primavera, cuando las temperaturas son altas, hace que muchas personas alérgicas al polen demanden servicios de urgencias.

Las condiciones meteorológicas adversas también provocan la proliferación de accidentes que requieren servicios de urgencia. Por lo tanto, está claro que para una mejor gestión de los recursos de emergencia, hay que tener en cuenta estos factores meteorológicos, ya que pueden influir en un aumento repentino de la demanda de asistencia de emergencia.

En este estudio se han tenido en cuenta algunos factores externos que pueden influir en el número de asistencias sanitarias, como los factores meteorológicos (por ejemplo, la temperatura mínima, máxima y media, las precipitaciones, la humedad y los datos diarios sobre la calidad del aire).

Los datos han sido descargados de la web de la Red de Información Ambiental de Andalucía REDIAM.

REDIAM tiene un total de 20 estaciones meteorológicas monitorizadas en toda la provincia de Jaén. Todos los datos atmosféricos recogidos en el sistema meteorológico son oficiales y corresponden al mismo periodo del conjunto de asistencias históricas tenidas en cuenta para nuestra investigación.

En concreto la estación con Id 15 está instalada en el nucleo urbano de Jaén y es de la que se han extraido todos los datos utilizados en nuestra investigación. En la figura 14 podemos ver un esquema de la composición final de la tabla.

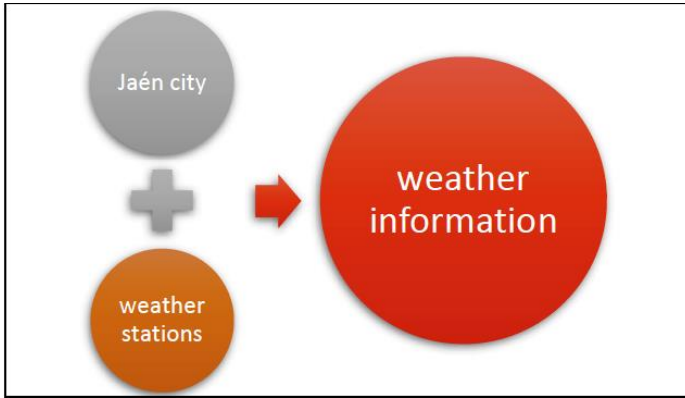


Figura 14 Composición definitiva de la tabla meteorológica.

3.1.1.3 Información sociológica

El núcleo urbano de Jaén se divide en nueve distritos denominados códigos postales. Cada asistencia este georeferenciada por el código postal, con lo cual se ha realizado una investigación exhaustiva de todos los datos sociológicos disponibles sobre esos códigos postales en organismos públicos y entidades privadas.

Del Instituto de Estadística y Cartografía de Andalucía se ha obtenido información sociológica distribuida en más de 100 campos, entre los más destacados encontramos:

- 1.- Número total de habitantes, por sexo y grupo de edad.
- 2.- Estado civil de la población, por grupo de edad.
- 3.- Nivel de estudios, por grupo de edad.
- 4.- Tipos de vivienda, uso, régimen, tamaño.
- 5.- Hogares y número de personas que lo componen.

3.1.2 Distribución en cuadrantes geolocalizados de Jaén.

Uno de los factores que enriquece la base de datos es la incorporación de información geolocalizada. Esto permite que la explotación de la base de datos tenga en cuenta la variabilidad de la información dependiendo de su ubicación.

La unidad geográfica mínima considerada es un cuadrante geolocalizado de 250 m, que es el utilizado por el Instituto de Estadística y Cartografía de Andalucía.

El mapa central de Jaén dividido en 128 cuadrantes con celdas de 250 m, lo podemos ver en la figura 15.

La información almacenada para cada cuadrante es:

- Identificación de cuadrantes.
- Coordenadas X, Y (UTM Zona 30, ETRS89) de las cuatro esquinas.
- Calles y número de ellas incluidas.
- Población total por grupos de edad.
- Información laboral: afiliados, por cuenta ajena, por cuenta propia y pensionados.
- Enlace con código postal.

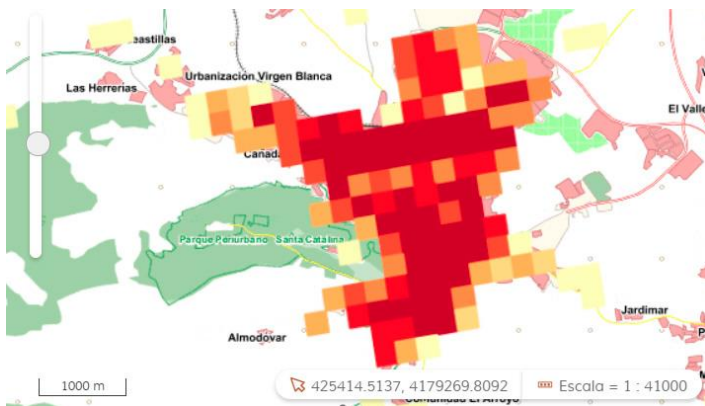


Figura 15 Cuadrantes geolocalizados que dividen la localidad de Jaén (Portal Web Instituto de estadística y cartografía de Andalucía).

3.1.3 Diagrama Entidad-Relacion

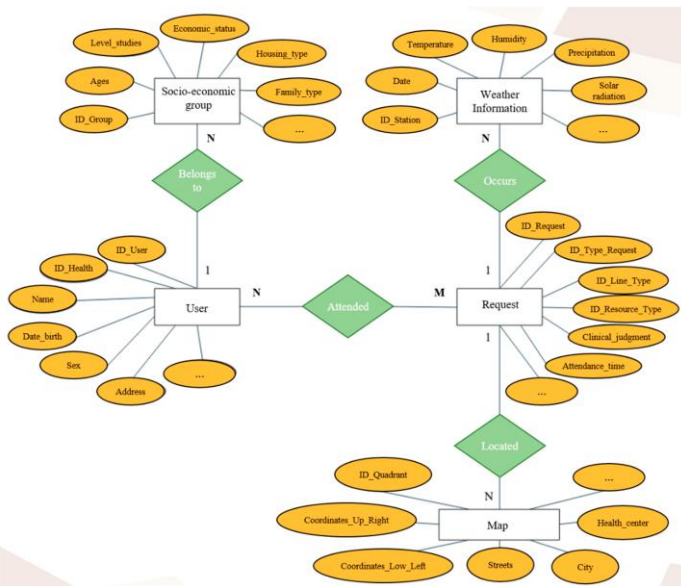


Figura 16 Diagrama Entidad Relación (DER).

El modelo entidad-relación de la figura 16 representa conceptualmente la organización y las relaciones de los datos en la base de datos diseñada.

En este caso, se trata de una representación simplificada, ya que la base de datos tiene múltiples tablas (más de 50 tablas). En el anexo están referenciadas las más significativas.

El objetivo de la figura 16 es mostrar el tipo de información almacenada y su relación. Cada entidad se agrupó como un bloque de información, y los bloques de datos representados fueron los siguientes:

- Datos socioeconómicos de los usuarios potencialmente atendibles.
- Información meteorológica y ambiental.
- Datos clínicos de los usuarios.
- Información geográfica de los usuarios.
- Recursos de emergencia movilizados.

El sistema gestor de base de datos utilizado fue el Oracle Database [68], que es un sistema de tipo objeto-relacional (ORD).

El entorno de desarrollo utilizado fue OracleSQL Developer [69], un entorno de desarrollo integrado que permite trabajar con SQL en bases de datos Oracle. Este entorno nos permitió crear y ejecutar consultas SQL y procedimientos para la integración de diferentes tipos de información.

3.1.4 Tablas, depuración de tablas y preparación de datos.

3.1.4.1 Tablas principales utilizadas

A continuación vamos a realizar un pequeño resumen de las tablas utilizadas a nivel práctico en nuestro proyecto. El

objetivo es simplemente tener una idea de que tipo de campos componen las tablas que utilizamos en cada uno de los pasos del desarrollo.

JAEN_V00_DEMANDAS

En esta tabla base JAEN_V00_DEMANDAS tenemos todos los recursos activados en todas las demandas sanitarias gestionadas en urgencias y emergencias en la ciudad de Jaén desde el año 2010 hasta febrero del 2019.

Esta tabla contiene la información de todo el proceso, desde que entra la llamada al centro de coordinación, hasta que finaliza la asistencia al paciente. Los detalles técnicos los podemos ver en la tabla 4.

JAEN_V00_DEMANDAS

Detalle tabla.

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
ANO_D	VARCHAR2(26 BYTE)	Yes	1
MES_D	VARCHAR2(6 BYTE)	Yes	2
DIA_D	VARCHAR2(6 BYTE)	Yes	3
IDPROVINCIA	VARCHAR2(26 BYTE)	Yes	4
FECHADEMANDA	VARCHAR2(26 BYTE)	Yes	5
IDDEMANDA	VARCHAR2(26 BYTE)	Yes	6
IDRECURSO	VARCHAR2(26 BYTE)	Yes	7
IDASISTENCIA	VARCHAR2(26 BYTE)	Yes	8
FECHALLAMADA	VARCHAR2(26 BYTE)	Yes	9
IDLLAMADA	VARCHAR2(26 BYTE)	Yes	10
IDLINEA	VARCHAR2(26 BYTE)	Yes	11
IDTIPOLINEA	VARCHAR2(26 BYTE)	Yes	12

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
IDALERTANTE	VARCHAR2(26 BYTE)	Yes	13
HORAENTRADA	VARCHAR2(26 BYTE)	Yes	14
HORACONTESTA	VARCHAR2(26 BYTE)	Yes	15
IDTIPODEMANDA	VARCHAR2(26 BYTE)	Yes	16
IDTIPODEMANDA1	VARCHAR2(26 BYTE)	Yes	17
IDTIPODEMANDA2	VARCHAR2(26 BYTE)	Yes	18
IDTIPODEMANDA3	VARCHAR2(26 BYTE)	Yes	19
FUENTECP	VARCHAR2(26 BYTE)	Yes	20
DIRIDPROVINCIARECURSO	VARCHAR2(256 BYTE)	Yes	21
DIRPEDANIARECURSO	VARCHAR2(256 BYTE)	Yes	22
DIRIDLOCALIDADRECURSO	VARCHAR2(256 BYTE)	Yes	23
DIRCALLE	VARCHAR2(256 BYTE)	Yes	24
DIRNUMERO	VARCHAR2(26 BYTE)	Yes	25
DIRPUERTA	VARCHAR2(26 BYTE)	Yes	26
DIRDT	VARCHAR2(26 BYTE)	Yes	27
DC	VARCHAR2(26 BYTE)	Yes	28
HORAACTIVACION	VARCHAR2(26 BYTE)	Yes	29
HORASALIDA	VARCHAR2(26 BYTE)	Yes	30
HORALLEGADA	VARCHAR2(26 BYTE)	Yes	31
HORACARGA	VARCHAR2(26 BYTE)	Yes	32
HORADESTINO	VARCHAR2(26 BYTE)	Yes	33
HORAOPERATIVA	VARCHAR2(26 BYTE)	Yes	34
HORADISPONIBLE	VARCHAR2(26 BYTE)	Yes	35
TIEMPOCOORDINACION	VARCHAR2(26 BYTE)	Yes	36
TIEMPOACTIVACION	VARCHAR2(26 BYTE)	Yes	37
TIEMPOATENCION	VARCHAR2(26 BYTE)	Yes	38
TIEMPORESPUESTA	VARCHAR2(26 BYTE)	Yes	39
IDMOTIVOEXCLUSION	VARCHAR2(26 BYTE)	Yes	40
IDUNIDADRECURSO	VARCHAR2(26 BYTE)	Yes	41
IDTIPORECURSO	VARCHAR2(26 BYTE)	Yes	42
IDEMPRESA	VARCHAR2(26 BYTE)	Yes	43
VEMERGENCIA	VARCHAR2(26 BYTE)	Yes	44

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
EDAD	VARCHAR2(26 BYTE)	Yes	45
SEXO	VARCHAR2(26 BYTE)	Yes	46
FNACIMIENTO	VARCHAR2(26 BYTE)	Yes	47
DIRIDPROVINCIA Paciente	VARCHAR2(256 BYTE)	Yes	48
DIRPEDANIA Paciente	VARCHAR2(256 BYTE)	Yes	49
DIRIDLOCALIDAD Paciente	VARCHAR2(256 BYTE)	Yes	50
DIRCALLE Paciente	VARCHAR2(256 BYTE)	Yes	51
DIRNUMERO Paciente	VARCHAR2(26 BYTE)	Yes	52
DIRPUERTA Paciente	VARCHAR2(26 BYTE)	Yes	53
JUICIO CLINICO1	VARCHAR2(26 BYTE)	Yes	54
JUICIO CLINICO2	VARCHAR2(26 BYTE)	Yes	55
JUICIO CLINICO3	VARCHAR2(26 BYTE)	Yes	56
ID CODIGO RESOLUCION	VARCHAR2(26 BYTE)	Yes	57
ID FINANCIACION	VARCHAR2(26 BYTE)	Yes	58
DIR DESTINODT	VARCHAR2(256 BYTE)	Yes	59
SITUACION	VARCHAR2(26 BYTE)	Yes	60
SITUACION DIRDT	VARCHAR2(10 BYTE)	Yes	61
ID CENTRO INGRESO	VARCHAR2(256 BYTE)	Yes	62
ID BASE	VARCHAR2(26 BYTE)	Yes	63
HORA LLAMADA	VARCHAR2(256 BYTE)	Yes	64
BASE	VARCHAR2(26 BYTE)	Yes	65
CENTRO	VARCHAR2(90 BYTE)	Yes	66
EMPRESA	VARCHAR2(26 BYTE)	Yes	67
HORA DC	VARCHAR2(256 BYTE)	Yes	68
PRIORIDAD	VARCHAR2(5 BYTE)	Yes	69
ORDEN	VARCHAR2(256 BYTE)	Yes	70
LTR	VARCHAR2(50 BYTE)	Yes	71
DIR ORIGENDT	VARCHAR2(256 BYTE)	Yes	72
PRIORIDAD TIPO RECURSO	VARCHAR2(5 BYTE)	Yes	73
ID TIPO RECURSO AGR	VARCHAR2(26 BYTE)	Yes	74
TIEMPO RECUPERACION	VARCHAR2(25 BYTE)	Yes	75
TIEMPO DEMORA	VARCHAR2(25 BYTE)	Yes	76

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
IDTIPOLINEAUNICO	VARCHAR2(23 BYTE)	Yes	77
HORACUELGA	VARCHAR2(256 BYTE)	Yes	78
CADREC	VARCHAR2(256 BYTE)	Yes	79
IDESCENARIO	VARCHAR2(26 BYTE)	Yes	80
TIPOEDAD	VARCHAR2(6 BYTE)	Yes	81
IDMOTIVOEXCLUSION2	VARCHAR2(26 BYTE)	Yes	82
HORACREACIONRECURSO	VARCHAR2(26 BYTE)	Yes	83
HORASUENA	VARCHAR2(256 BYTE)	Yes	84
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	85
DIA_SEMANA	VARCHAR2(10 BYTE)	Yes	86

Tabla 4 JAEN_V00_DEMANDAS.

JAEN_EX_INF_METEROROLOGICA

En esta tabla almacenamos toda la información meteorológica desde el año 2010 hasta febrero de 2019.

Dicha información será enlazada con la tabla anterior para poder estudiar como influyen todos estos parámetros meteorológicos en la actividad sanitaria diaria de urgencias y emergencias de la ciudad de Jaén. Tabla 5.

JAEN_EX_INF_METEOROLOGICA

Detalle tabla

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
PROVINCIA_ID	VARCHAR2(2 BYTE)	No	1
ESTACION_ID	NUMBER	No	2
FECHA	DATE	No	3
DIA_JULIANO	NUMBER	Yes	4

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
T_MAX	NUMBER	Yes	5
HT_MAX	VARCHAR2(5 CHAR)	Yes	6
T_MIN	NUMBER	Yes	7
HT_MIN	VARCHAR2(5 CHAR)	Yes	8
T_MED	NUMBER	Yes	9
HUM_MAX	NUMBER	Yes	10
HUM_MIN	NUMBER	Yes	11
HUM_MED	NUMBER	Yes	12
VEL_VIENTO	NUMBER	Yes	13
DIR_VIENTO	NUMBER	Yes	14
RADIACION	NUMBER	Yes	15
PRECIPITACION	NUMBER	Yes	16
ETO	NUMBER	Yes	17

Tabla 5 JAEN_EX_INF_METEOROLOGICA.

JAEN_EX_PROVINCIA

Al obtener los datos de las estaciones meteorológicas oficiales de la Junta de Andalucía, disponemos de información de todas las provincias. Tabla 6.

En nuestro caso nos quedamos sólo con la información de la provincia de Jaén y más concretamente de la estación 15 que es la que cubre la ciudad de Jaén.

JAEN_EX_PROVINCIA

Detalle tabla

COLUM_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
PROVINCIA_ID	VARCHAR2(2 BYTE)	No	1
PROVINCIA	VARCHAR2(100 BYTE)	No	2

Tabla 6 JAEN_EX_PROVINCIA.

JAEN_EX_ESTACION_METEOROLOGICA

Datos del conjunto de estaciones meteorológicas que componen la provovincia de Jaén, como indicamos con anterioridad la estación 15 es la que cubre la ciudad de Jaén. Detalles técnicos en la siguiente tabla 7.

JAEN_EX_ESTACION_METEOROLOGICA

Detalle tabla

COLUM_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
PROVINCIA_ID	VARCHAR2(2 BYTE)	No	1
ESTACION_ID	NUMBER(5,0)	No	2
ESTACION	VARCHAR2(100 BYTE)	No	3

Tabla 7 JAEN_EX_ESTACION_METEOROLOGICA.

JAEN_V00_CUADRANTE_JAEN

La ciudad de Jaén la hemos dividido en 128 Cuadrantes de 250m x 250m cada uno. De cada uno de estos cuadrantes hemos almacenado las coordenadas UTM de los cuatro vértices para tenerlo geolocalizado.

De cada uno de estos cuadrantes hemos obtenido información de diverso tipo como son: Calles, números de la calle que están incluidos, Código postal, población, centro de salud a la que pertenecen dichas calles, Grupos de edad, ocupación e información general socioeconómica de dicha población. Tabla 8.

JAEN_V00_CUADRANTE_JAEN

Detalle tabla

COLUM_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
CUADRANTE	NUMBER(5,0)	Yes	1
X_SI	NUMBER(19,11)	Yes	2
Y_SI	NUMBER(17,10)	Yes	3
X_SD	NUMBER(17,11)	Yes	4
Y_SD	NUMBER(19,10)	Yes	5
X_II	NUMBER(19,11)	Yes	6
Y_II	NUMBER(19,10)	Yes	7
X_ID	NUMBER(19,11)	Yes	8
Y_ID	NUMBER(19,10)	Yes	9
DIRECCIÓN1	VARCHAR2(512 BYTE)	Yes	10
NUMERO_DESDE	NUMBER(26,0)	Yes	11
NUMERO_HASTA	NUMBER(26,0)	Yes	12
POBLACION_TOTAL	NUMBER(10,0)	Yes	13
GRUPOS_EDAD_MENOR_16	NUMBER(10,0)	Yes	14
GRUPOS_EDAD_DE_16_A_64	NUMBER(10,0)	Yes	15
GRUPOS_EDAD_65_Y_MAS	NUMBER(10,0)	Yes	16
AFILIACION_SS	NUMBER(10,0)	Yes	17
T_CUENTA_PROPIA	NUMBER(10,0)	Yes	18
T_CUENTA_AJENA	NUMBER(10,0)	Yes	19
PENSIONISTAS	NUMBER(10,0)	Yes	20
CP	NUMBER(10,0)	Yes	21

COLUM_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
CENTRO_DE_SALUD	VARCHAR2(26 BYTE)	Yes	22
DIRECCIÓN2	VARCHAR2(512 BYTE)	Yes	23

Tabla 8 JAEN_V00_CUADRANTE_JAEN.

JAEN_V00_CALLES_CP

Esta tabla contiene todas las calles de la ciudad de Jaén con su numeración y a que código postal pertenece la numeración. Tabla 9.

JAEN_V00_CALLES_CP

Detalle tabla

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
TIPO_VIA	VARCHAR2(128 BYTE)	Yes	1
NOMBRE_VIA	VARCHAR2(512 BYTE)	Yes	2
NUMERACION	VARCHAR2(128 BYTE)	Yes	3
EXTREMO_INFERIOR	NUMBER(26,0)	Yes	4
EXTREMO_SUPERIOR	NUMBER(26,0)	Yes	5
CP	NUMBER(26,0)	Yes	6

Tabla 9 JAEN_V00_CALLES_CP.

JAEN_V00_VALORES_X_CP

En este tabla hemos recogido la máxima información posible sobre las características sociales, económicas, educativas y de vivienda de cada código postal. Tabla 10.

JAEN_V00_VALORES_X_CP

Campos de la tabla

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
CPRO	NUMBER(4,0)	Yes	1
CMUN	NUMBER(2,0)	Yes	2
CP	NUMBER(5,0)	Yes	3
POBLACION_TOTAL	NUMBER(5,0)	Yes	4
HOMBRES	NUMBER(5,0)	Yes	5
MUJERES	NUMBER(5,0)	Yes	6
PERSONAS_MENOS_16	NUMBER(4,0)	Yes	7
PERSONAS_16_A_64	NUMBER(5,0)	Yes	8
PERSONAS_MAS_64	NUMBER(4,0)	Yes	9
HOMBRES_MENOS_16	NUMBER(4,0)	Yes	10
HOMBRES_16_A_64	NUMBER(5,0)	Yes	11
HOMBRES_MAS_64	NUMBER(4,0)	Yes	12
MUJERES_MENOS_16	NUMBER(4,0)	Yes	13
MUJERES_16_A_64	NUMBER(5,0)	Yes	14
MUJERES_MAS_64	NUMBER(4,0)	Yes	15
PERSONAS_SOLTERAS	NUMBER(5,0)	Yes	16
PERSONAS_CASADAS	NUMBER(5,0)	Yes	17
PERSONAS_SEPARADAS	NUMBER(3,0)	Yes	18
PERSONAS_DIVORCIADAS	NUMBER(3,0)	Yes	19
PERSONAS_VIUDAS	NUMBER(4,0)	Yes	20
PERSONAS_SOLTERAS_MENOS_16	NUMBER(4,0)	Yes	21
PERSONAS_SOLTERAS_16_A_64	NUMBER(4,0)	Yes	22
PERSONAS_SOLTERAS_MAS_64	NUMBER(3,0)	Yes	23
PERSONAS_CASADAS_MENOS_16	NUMBER(2,0)	Yes	24
PERSONAS_CASADAS_16_A_64	NUMBER(5,0)	Yes	25
PERSONAS_CASADAS_MAS_64	NUMBER(4,0)	Yes	26
PERSONAS_SEPARADAS_MENOS_16	NUMBER(1,0)	Yes	27
PERSONAS_SEPARADAS_16_A_64	NUMBER(3,0)	Yes	28

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
PERSONAS_SEPARADAS_MAS_64	NUMBER(2,0)	Yes	29
PERSONAS_DIVORCIADAS_MENOS_16	NUMBER(1,0)	Yes	30
PERSONAS_DIVORCIADAS_16_Y_64	NUMBER(3,0)	Yes	31
PERSONAS_DIVORCIADAS_MAS_64	NUMBER(2,0)	Yes	32
PERSONAS_VIUDAS_MENOS_16	NUMBER(2,0)	Yes	33
PERSONAS_VIUDAS_ENTRE_16_Y_64	NUMBER(3,0)	Yes	34
PERSONAS_VIUDAS_MAS_64	NUMBER(4,0)	Yes	35
PERSONAS_ANALFABETAS	NUMBER(3,0)	Yes	36
PERSONAS_SIN_ESTUDIOS	NUMBER(4,0)	Yes	37
PER_ESTUDIOS_PRIMER_GRADO	NUMBER(4,0)	Yes	38
PER_CON_ESTUDIOS_DE_2_GRADO	NUMBER(5,0)	Yes	39
PER_CON_ESTUDIOS_DE_3_GRADO	NUMBER(4,0)	Yes	40
SIN_DATOS_DE_ESTUD_MENOS_16	NUMBER(4,0)	Yes	41
HOMBRES_ANALFABETOS	NUMBER(2,0)	Yes	42
MUJERES_ANALFABETAS	NUMBER(3,0)	Yes	43
HOMBRES_SIN_ESTUDIOS	NUMBER(3,0)	Yes	44
MUJERES_SIN_ESTUDIOS	NUMBER(4,0)	Yes	45
HOMBRES_ESTUDIOS_PRIMER_GRADO	NUMBER(4,0)	Yes	46
MUJERES_ESTUDIOS_PRIMER_GRADO	NUMBER(4,0)	Yes	47
HOMBRES_ESTUDIOS_2_GRADO	NUMBER(4,0)	Yes	48
MUJERES_ESTUDIOS_2_GRADO	NUMBER(4,0)	Yes	49
HOMBRES_ESTUDIOS_TERCER_GRADO	NUMBER(4,0)	Yes	50
MUJERES_ESTUDIOS_TERCER_GRADO	NUMBER(4,0)	Yes	51
H_SIN_DATOS_DE_ESTUD_MEN_16	NUMBER(4,0)	Yes	52
M_SIN_DATOS_DE_ESTUD_MEN_16	NUMBER(4,0)	Yes	53
PERSONAS_16_Y_64_ANALFAB	NUMBER(3,0)	Yes	54
PERSONAS_MAS_64_ANALFABETAS	NUMBER(3,0)	Yes	55
PERSONAS_16_Y_64_SIN_ESTUDIOS	NUMBER(3,0)	Yes	56
PER_MAS_64_SIN_ESTUDIOS	NUMBER(4,0)	Yes	57
PER_16_Y_64_ESTUD_1ER_G	NUMBER(4,0)	Yes	58
PER_MAS_64_ESTUD_1ER_G	NUMBER(3,0)	Yes	59
PER_16_Y_64_ESTUD_2_G	NUMBER(5,0)	Yes	60

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
PER_MAS_64_ESTUD_2_G	NUMBER(4,0)	Yes	61
PER_16_Y_64_ESTUD_3ER_G	NUMBER(4,0)	Yes	62
PER_MAS_64_ESTUD_3ER_G	NUMBER(3,0)	Yes	63
PER_MENOS_16_SIN_INF_ESTUD	NUMBER(4,0)	Yes	64
TOTAL_VIVIENDAS	NUMBER(5,0)	Yes	65
VIVIENDAS_PRINCIPALES	NUMBER(5,0)	Yes	66
VIVIENDAS_SECUNDARIAS	NUMBER(4,0)	Yes	67
VIVIENDAS_VACIAS	NUMBER(4,0)	Yes	68
VIV_EN_PROP_POR_COMPRA_PAG	NUMBER(4,0)	Yes	69
VIV_EN_PROP_PAGOS_PEND	NUMBER(4,0)	Yes	70
VIV_EN_PROP_HERENCIA_O_DONA	NUMBER(3,0)	Yes	71
VIV_EN_ALQUILER	NUMBER(3,0)	Yes	72
VIV_CEDIDA_O_BAJO_PRECIO	NUMBER(3,0)	Yes	73
VIV_EN_OTRO_TIPO_DE_TENENCIA	NUMBER(3,0)	Yes	74
VIV_DE_MENOS_DE_30M2	NUMBER(1,0)	Yes	75
VIV_30_45_M2	NUMBER(4,0)	Yes	76
VIV_46_60_M2	NUMBER(6,0)	Yes	77
VIV_61_75_M2	NUMBER(6,0)	Yes	78
VIV_76_90_M2	NUMBER(6,0)	Yes	79
VIV_91_105_M2	NUMBER(6,0)	Yes	80
VIV_106_120_M2	NUMBER(5,0)	Yes	81
VIV_121_150_M2	NUMBER(5,0)	Yes	82
VIV_151_180_M2	NUMBER(5,0)	Yes	83
VIV_DE_MAS_DE_180_M2	NUMBER(5,0)	Yes	84
VIV_CON_1_HABITACION	NUMBER(2,0)	Yes	85
VIV_CON_2_HAB	NUMBER(2,0)	Yes	86
VIV_CON_3_HAB	NUMBER(3,0)	Yes	87
VIV_CON_4_HAB	NUMBER(4,0)	Yes	88
VIV_CON_5_HAB	NUMBER(4,0)	Yes	89
VIV_CON_6_HAB	NUMBER(4,0)	Yes	90
VIV_CON_7_HAB	NUMBER(3,0)	Yes	91
VIV_CON_8_HAB	NUMBER(2,0)	Yes	92

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
VIV_CON_9_O_MAS_HAB	NUMBER(2,0)	Yes	93
TOTAL_HOG	NUMBER(5,0)	Yes	94
HOG_DE_1_PERSONA	NUMBER(4,0)	Yes	95
HOG_DE_2_PER	NUMBER(4,0)	Yes	96
HOG_DE_3_PER	NUMBER(4,0)	Yes	97
HOG_DE_4_PER	NUMBER(4,0)	Yes	98
HOG_DE_5_PER	NUMBER(3,0)	Yes	99
HOG_DE_6_O_MAS_PER	NUMBER(3,0)	Yes	100

Tabla 10 JAEN_V00_VALORES_X_CP.

JAEN_V01_CALLES_DIRDT

En esta tabla tenemos relacionadas todas las viviendas de la ciudad de Jaén con el centro de salud al que pertenece. En la tabla 11 vemos el detalle de campos.

Detalle tabla

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
CALLE	VARCHAR2(256 BYTE)	Yes	1
NUMERO	VARCHAR2(26 BYTE)	Yes	2
CENTRO_SALUD	VARCHAR2(26 BYTE)	Yes	3

Tabla 11 JAEN_V01_CALLES_DIRDT.

3.1.4.2 Depuración de datos.

Comenzamos una labor de explotación para dar utilidad a dicha estructura. Después de estudiar todas las posibilidades

que nos da nuestra base de datos, nos fijamos un objetivo principal, crear un sistema capaz de predecir la necesidad de recursos sanitarios en distintas situaciones de la actividad de urgencias y emergencias de la ciudad de Jaén. Para este primer objetivo ponemos el foco en las patologías respiratorias.

Comenzamos un proceso de investigación para conseguir un entorno de base de datos adecuado para afrontar este primer reto.

El primer paso es la adaptación de la tabla JAEN_V00_DEMANDAS para crear una nueva tabla base en las que tengamos sólo la información relativa a las patologías respiratorias.

A lo largo del proceso de realización de ese sistema veremos como es necesario añadir nuevos campos calculados a esta tabla para conseguir ajustar con mayor precisión nuestro sistema.

La primera vía de explotación la hemos realizado tomando las demandas con patologías respiratorias.

Hemos creado la tabla JAEN_V02_DEMANDAS_AMP con todos los recursos activados para todas las demandas con dicha patología.

Una vez seleccionados los recursos activados para esta patología, hemos realizado una serie de depuraciones para asegurarnos que toda la información de dicha tabla es válida para que nuestro sistema de predicción la utilice.

A continuación explicamos con detalle cada proceso de depuración. Estos procesos de depuración son necesarios para eliminar datos anómalos de nuestro sistema que podrían llevar a error al sistema predictivo, ya que por defecto o exceso de

información pueden crear tendencias erróneas en la movilización de recursos que no se corresponden con la realidad.

3.1.4.2.1 Primer proceso de depuración.

En la tabla JAEN_V02_DEMANDAS_AMP teníamos 103.790 registros, en estos registros, además de toda la información válida de recursos, nos encontramos registros que pertenecen a demandas en las que no ha sido necesario activar ningún recurso, es decir, se le ha dado solución mediante un consejo sanitario.

Los registros con código de resolución (4*) son aquellas demandas resueltas con un consejo sanitario y cuyo campo idrecurso está a null. En un primer barrido de la tabla borramos 1.332 registros que tienen a null el idrecurso.

En un segundo barrido buscamos los códigos de resolución (00) que son las demandas en las que el usuario rechaza el ofrecimiento del recurso.

Después de estos dos filtrados nos quedan un total de 102458 registros.

3.1.4.2.2 Segundo proceso de depuración

Al tener información de más de 9 años, tenemos que buscar y eliminar todos aquellos días en los que por motivos técnicos o de otra índole falte información a lo largo de las 24 horas de dicho día.

Tenemos que asegurarnos que cuando un día lo tenemos en cuenta para todos los cálculos de nuestro sistema es porque el día es válido y no nos mete ruido en todos los procesos, incluido el proceso final de predicción.

Estudiamos y depuramos todos los días de todos los años en los que detectemos que ha habido un problema técnico que haya hecho que dicho día no tenga toda la información completa, así evitamos resultados anómalos que nos desvirtuen la fiabilidad del sistema.

Analizando la totalidad de los días, estos son los días eliminados de la base de datos en cada uno de los años.

Detalle de días eliminados en cada año afectado.

AÑO 2010

Falta el día 30/09/2010...Se elimina de la tablas auxiliares cono son las tablas de días festivos y laborales.

AÑO 2011

Todos los días son correctos y por tanto válidos para nuestro sistema.

AÑO 2012

Falta 09/04/2012

Realizamos una consulta para ver el número de demandas totales de cualquier patología que tenemos cada día.

En base a todas las comprobaciones y análisis establecemos que un día tiene que tener al menos 350 demandas de cualquier tipo a lo largo de las 24h para que pueda ser válido. En todos los días con un menor número de demandas hemos comprobado que hay intervalos horarios con ausencia de información.

Año 2012

En la tabla 12 vemos los días con número de demandas totales inferior a 350.

FECHADEMANDA	COUNT(FECHADEMANDA)
20120606	16
20120626	26
20120628	70
20120411	82
20120624	102
20120607	116
20120421	128
20120505	144
20120610	164
20120402	184
20120619	204

FECHADEMANDA	COUNT(FECHADEMANDA)
20120620	216
20120115	217
20120604	218
20120419	236
20121224	240
20120611	242
20120615	252
20121120	269
20120601	270
20120605	272
20120531	274
FECHADEMANDA	COUNT(FECHADEMANDA)
20121217	284
20120520	284
20121119	301
20120407	308
20121216	314
20120522	314
20120428	322
20121127	323
20120117	324

Tabla 12 Días con incidencia en el año 2012.

AÑO 2013

Todos los días son correctos y por tanto válidos para nuestro sistema.

AÑO 2014

Todos los días son correctos y por tanto válidos para nuestro sistema.

AÑO 2015

Días con número de demandas totales inferior a 350. Tabla 13.

FECHADEMANDA	COUNT(FECHADEMANDA)
20150714	198
20150410	236
20150430	251
20150611	253
20150712	258
20150715	260
20150506	265
20150918	267
20150922	272
20150919	273
20150423	274

FECHADEMANDA	COUNT(FECHADEMANDA)
20150701	278
20150921	279
20150713	283
20150716	289
20150730	292
20150522	292
20150920	292
20150417	292

20150819	293
20150917	293
20150804	293

FECHADEMANDA	COUNT(FECHADEMANDA)
20150702	293
20150805	294
20150814	301
20150916	301
20150824	302
20150901	304
20150420	304
20150710	304
20150503	310
20150509	310
20150707	311

FECHADEMANDA	COUNT(FECHADEMANDA)
20150724	312
20150923	314
20150614	316
20151008	316
20150813	318
20150519	318
20151001	319
20151119	319
20151222	320
20150818	320
20150524	320

FECHADEMANDA	COUNT(FECHADEMANDA)
20150426	321
20150711	322
20150820	327
20150825	328
20151118	330

Tabla 13 Días con incidencia en el año 2015

AÑO 2016

Todos los días son correctos y por tanto válidos para nuestro sistema.

Año 2017

FALTA 28/06/2017
 29/09/2017

Año 2018

Días con número de demandas totales inferior a 350.

FECHADEMANDA	COUNT(FECHADEMANDA)
20180117	11

La versión definitiva tabla JAEN_V02_DEMANDAS_AMP después de pasar todos los procesos de depuración tiene 96.711 registros.

La estructura explicada anteriormente tiene muchas aplicaciones, una de las más importantes es estudiar el número de asistencias que se esperan por demarcación, teniendo en cuenta las distintas variables que afectan al resultado.

Profundizando en este supuesto, tenemos que generar nuevas tablas intermedias en las que se almacenará la información agrupada por días, zonas y recursos.

Obtendremos así una previsión del número y tipo de asistencias esperadas, así como una valiosa información sobre la movilización de recursos que se espera en las diferentes épocas del año.

Para este objetivo tendremos en cuenta además variables como los datos atmosféricos, el día de la semana, festivo o laboral, los valores económicos de la zona y todas las variables que rodean a una asistencia.

La primera tabla intermedia construida contiene el número de recursos activados en cada semana, distinguiendo si el tipo de día es festivo o laboral.

Tiene la siguiente estructura:

Tabla JAEN_V11_DEMANDA_FES_AG_CTRL. Nueva tabla intermedia generada para la optimización de la gestión de recursos. Detalle de campos en la tabla 14.

JAEN_V11_DEMANDA_FES_AG_CTRL

Detalle tabla

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TEMPERATURA_MAXIMA	NUMBER	Yes	4
TEMPERATURA_MINIMA	NUMBER	Yes	5
TEMPERATURA_MEDIA	NUMBER	Yes	6
HUMEDAD_MAXIMA	NUMBER	Yes	7
TEMPERATURA_MINIMA	NUMBER	Yes	8
TEMPERATURA_MEDIA	NUMBER	Yes	9
VELOCIDAD_VIENTO	NUMBER	Yes	10
RADIACION	NUMBER	Yes	11
PRECIPITACION	NUMBER	Yes	12
ETO	NUMBER	Yes	13
N_RECURSOS	NUMBER	Yes	14

Tabla 14 JAEN_V11_DEMANDA_FES_AG_CTRL.

Finalmente creamos la última tabla en la que calculamos la media de recursos activados cada semana.

Añadimos el campo TIPO_DIA con posibles valores Festivo y Laboral.

Tabla JAEN_V12_DEM_AMP_FES_AG_CTRL
Tabla final sobre la que aplicaremos los modelos predictivos.
Tabla 15.

JAEN_V12_DEM_AMP_FES_AG_CTRL

Detalle tabla

COLUM_NAME	DATA_TYPE	NULLABLE	C_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	4
TEMPERATURA_MAXIMA	NUMBER	Yes	5
TEMPERATURA_MINIMA	NUMBER	Yes	6
TEMPERATURA_MEDIA	NUMBER	Yes	7
HUMEDAD_MAXIMA	NUMBER	Yes	8
TEMPERATURA_MINIMA	NUMBER	Yes	9
TEMPERATURA_MEDIA	NUMBER	Yes	10
VELOCIDAD_VIENTO	NUMBER	Yes	11
RADIACION	NUMBER	Yes	12
PRECIPITACION	NUMBER	Yes	13
ETO	NUMBER	Yes	14
N_RECURSOS	NUMBER	Yes	15
N_DIAS	NUMBER	Yes	16
MEDIA_RECURSOS	NUMBER	Yes	17

Tabla 15 JAEN_V12_DEM_AMP_FES_AG_CTRL.

El crecimiento de la base de datos y el aumento del volumen de información almacenada significa que ahora se dispone de valiosos datos históricos.

La futura explotación y análisis de estos datos, como la detección de patrones de comportamiento y relaciones entre variables, permitirá planificar con antelación los recursos de emergencia disponibles en cada momento del año y en cada parte de la ciudad.

3.2 Modelo Predictivo de recursos sanitarios en casos de dificultad respiratoria o Disnea.

3.2.1 Estudio y análisis del estado actual

Siguiendo el flujo de trabajo de la metodología descrita anteriormente, se analizaron los datos iniciales.

En este caso, se trata de los datos sobre el número de recursos movilizados para atender las demandas de urgencias por patologías respiratorias. Se han realizado varios análisis a diferentes niveles de agrupación.

En primer lugar, se han agrupado los datos por meses, tal y como se muestra en la tabla 16.

Year	January	February	March	April	May	June	July	August	September	October	November	December
2012	222	165	206	150	153	127	88	109	89	160	129	132
2013	163	144	189	190	131	119	112	120	99	99	129	162
2014	188	292	281	200	246	98	116	101	121	144	169	139
2015	214	158	196	213	199	184	126	101	120	145	170	193
2016	357	192	188	202	171	132	88	100	158	152	168	237
2017	370	240	233	167	129	99	37	53	43	76	99	81
2018	154	154	176	133	142	127	123	85	115	97	143	124
2019	233	160	122	133	143	100	121	93	112	101	131	193

Tabla 16 Recursos sanitarios de urgencia activados para atender a pacientes con enfermedades respiratorias, agrupados por meses de 2012 a 2019.

A continuación se ha analizado la evolución de estos datos en el tiempo, figura 17

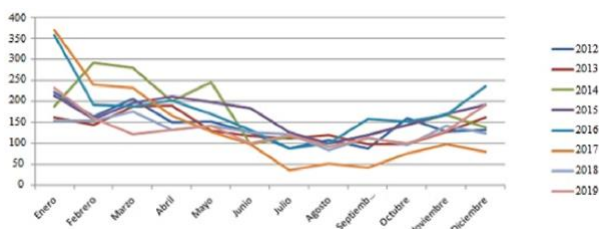


Figura 17 Distribución de los recursos sanitarios urgentes activados para atender a pacientes con enfermedades respiratorias, agrupados por meses desde 2012 hasta 2019.

En esta última se observa que en los primeros meses del año y en los últimos, correspondientes a otoño e invierno, el número de recursos movilizados es mayor que en los meses con buenas condiciones climatológicas, como primavera y sobre todo verano.

Al analizar en detalle las series de la figura 17, se observa un comportamiento anómalo en la tendencia durante 2014 en los meses de febrero a mayo, ya que se observa un incremento muy acusado en la activación de recursos sanitarios.

Este tipo de situaciones apoyan la necesidad de utilizar la minería de datos para obtener modelos capaces de predecir este tipo de situaciones. Ya que, si seguimos un modelo pasivo, es decir, promediando, las decisiones tomadas para dimensionar los recursos sanitarios en estos meses de 2014 no fueron las adecuadas para prestar un servicio de emergencias de calidad.

En el caso de utilizar el valor promediado no estaríamos teniendo en cuenta el diferente comportamiento de la demanda de recursos que existe a lo largo del año. Por tanto, nos encontraríamos con meses en los que los recursos estén sobredimensionados y otros periodos en los que falten recursos para dar respuesta a todas las demandas de los usuarios.

Utilizando nuestro sistema, conseguimos un modelo dinámico que tenga en cuenta los periodos de mayor y menor demanda y se ajunte a dicha tendencia.

En el estudio inicial también se analizó la distribución de las demandas de asistencia a lo largo de la semana, tabla 17. Creemos que dicha distribución es necesaria ya que entendemos que a lo largo de la semana existe razones por las que la demanda de recursos puede variar entre unos días y otros. Como veremos con detalle posteriormente, dependiendo del día de la semana, los usuarios van a tener disponibles otro servicios sanitarios distintos al de urgencias y emergencias.

Year	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
2012	274	242	212	194	268	261	279
2013	260	220	199	229	225	257	267
2014	257	277	288	307	335	277	354
2015	279	270	271	270	279	350	300
2016	283	279	307	286	328	299	363
2017	266	213	210	193	210	245	290
2018	225	185	212	221	223	254	253
2019	253	222	191	237	229	256	254

Tabla 17 Recursos sanitarios de urgencias y emergencias activados para la atención de pacientes con enfermedades respiratorias, agrupados por días de la semana de 2012 a 2019.

Para una mejor visualización de estos datos, mostramos dicha tendencia en la siguiente fig.18. Posteriormente analizaremos con más detalle el comportamiento.

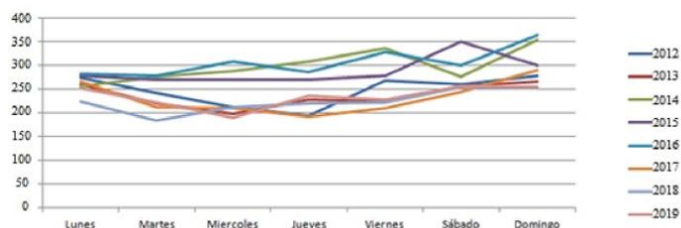


Figura 18 Distribución de los recursos sanitarios urgentes activados para atender a pacientes con enfermedades respiratorias, agrupados por días de la semana desde 2012 hasta 2019.

En la figura 18 se observa que la distribución de las demandas aumenta los fines de semana. Esto se debe a que entre semana los pacientes disponen de otras alternativas asistenciales, como acudir a los centros de atención primaria, y no tienen que recurrir al servicio de urgencias.

Aunque la diferencia es apreciable, no puede ser excesivamente significativa, ya que no debemos olvidar que el teléfono de urgencias está destinado a pacientes con patología y síntomas importantes, en cuyos casos los pacientes acuden directamente a los servicios de urgencias.

Al analizar las demandas por días de la semana, se observa que los lunes hay un ligero aumento de las demandas respecto al resto de días laborables. Esto puede deberse a que los sábados y domingos los centros de salud de atención primaria no abren, y estos pacientes no han podido ser atendidos por su médico durante el fin de semana.

Por consiguiente, la demanda es menor los demás días, ya que los centros de atención primaria están abiertos entre semana y muchos pacientes pueden ser atendidos por los médicos de estos centros. Así, también se ha observado que la evolución de estos datos en días festivos es similar a la de los fines de semana.

Teniendo en cuenta toda esta información y tras observar que existe una gran diferencia en el comportamiento de las demandas urgentes en días laborables y festivos, en este estudio predictivo se tratarán estos datos por separado, generando modelos diferentes. Un modelo de predicción para los días laborables, es decir, entre semana, y otro modelo de predicción para los días festivos, es decir, los fines de semana y festivos.

Dicho esto, para la construcción de estos modelos se ha estimado que la mejor agrupación de los datos es por: Año, mes, semana, tipo de día (laborable o fin de semana/festivo).

La tabla 18 muestra un ejemplo de la tabla que se utilizará para elaborar uno de los modelos, con los datos que servirán de entrada para generar el modelo predictivo. En ella se muestra el nivel de agrupación por días, distinguiendo entre día laborable o fin de semana/festivo.

En este estudio se utiliza un ciclo de vida clásico de minería de datos: Una vez aplicados los algoritmos de limpieza para detectar anomalías en los datos y agrupados éstos, se prueban varios modelos de regresión.

Los resultados de cada modelo se analizan estadísticamente mediante validación cruzada para identificar el más adecuado.

Year	Month	Week	Type of day	Max Temp.	Min Temp.	Mean Temp.	Max Hum.	Min Hum.	Mean Hum.	Wind speed	Rad.	Rainfall	Num Resources	Num Days	Mean Res.
2019	1	1	Weekday	14.6	1.6	7.7	95.8	49.5	80.8	0.5	6.1	0.0	23	3	7.7
2019	1	2	Weekday	15.3	-1.3	5.8	97.3	47.4	81.0	0.6	9.0	0.1	40	5	8.0
2019	1	3	Weekday	9.5	-1.3	3.8	91.0	40.1	70.4	0.9	7.3	0.7	33	5	6.6
2019	1	4	Weekday	13.5	-1.0	5.4	96.3	42.1	77.3	0.9	9.9	1.6	25	5	5.0
2019	2	1	Weekday	17.0	4.4	10.1	93.1	47.6	76.2	1.1	9.5	0.1	30	5	6.0
2019	2	2	Weekday	15.3	1.0	7.8	95.7	40.1	76.9	0.9	11.1	3.7	23	5	4.6
2019	2	3	Weekday	18.8	3.9	11.0	90.9	43.5	65.7	1.0	12.8	0.0	38	5	7.6
2019	2	4	Weekday	18.9	7.6	13.5	92.5	39.5	67.9	1.0	9.1	1.8	15	4	3.8

Tabla 18 Tabla de ejemplo del conjunto de datos utilizado para generar los modelos predictivos.

3.2.2 Técnicas de minería de datos utilizadas

La minería de datos es una combinación de estadística, informática y ML (Machine Learning).

Las técnicas utilizadas en la minería de datos implican un conjunto de cálculos para crear un modelo a partir de un conjunto de datos seleccionados.

En este sentido, para crear dicho modelo, un algoritmo analiza primero los datos proporcionados, buscando tipos específicos de patrones o tendencias. Utiliza los resultados de este análisis a lo largo de muchas iteraciones para encontrar los parámetros óptimos para generar el modelo.

A continuación, estos parámetros se aplican a todo el conjunto de datos para extraer patrones procesables y estadísticas detalladas.

En este trabajo, el análisis completo de los datos se ha llevado a cabo con un módulo de análisis incluido en el software Oracle DataMining [70]. En esta fase, el reto consiste en elegir el algoritmo que mejor realice la predicción del atributo objetivo.

Una vez que se dispone del conjunto de datos, hay que llevar a cabo la detección de anomalías en los datos antes de que pasen a formar parte de los datos de entrada.

La detección de anomalías es importante para detectar fraudes, valores atípicos y otros sucesos raros que pueden tener un gran impacto negativo en el proceso global y que, a priori, son difíciles de detectar.

En esta fase se utiliza un algoritmo de clasificación, ya que la detección de anomalías puede considerarse un tipo de clasificación.

La desviación del perfil se identifica como una anomalía. En concreto, en esta fase, el algoritmo utilizado ha sido el de la máquina de vectores de soporte (SVM) [64, 71, 72].

Se trata de un algoritmo que se ha utilizado recientemente en trabajos de predicción relacionados con datos médicos [73-78].

SVM trabaja con la idea básica de minimizar la hiperesfera de la clase única de ejemplos en los datos de entrenamiento y considera todas las demás muestras fuera de esta hiperesfera como valores atípicos o fuera de la distribución de los datos de entrenamiento.

Este algoritmo produce una predicción y una probabilidad para cada caso de los datos. Si la predicción es 1, el caso se considera típico. Si la predicción es 0, el caso se considera anómalo.

Siguiendo el flujo de trabajo ya explicado en las secciones anteriores, antes de crear el modelo se calculó el nivel de influencia que cada una de las variables de entrada tiene sobre el atributo objetivo utilizando el algoritmo de Longitud de Descripción Mínima (MDL) [64].

La ejecución de este algoritmo devuelve un valor entre - 1 y 1. El valor - 1 significa que esta variable no tiene relación con el atributo objetivo y que puede introducir ruido en el modelo, por lo que debe descartarse del cálculo del modelo. El valor 0 significa que la variable no está relacionada con el objetivo y 1 significa una relación máxima entre la variable y el atributo objetivo.

El algoritmo MDL trata cada atributo como un modelo predictivo de la clase objetivo. Cada modelo predictivo de estos atributos se compara y clasifica en la métrica MDL. MDL penaliza la complejidad del modelo para evitar el sobreajuste.

En este sentido, antes de introducir los datos en los algoritmos, como se indica en la sección 2.3, se realizó una exploración y un análisis de los datos mediante técnicas de distribución, histogramas y detección de anomalías. De este modo se evitó tanto el sobreajuste como el infraajuste de los modelos al aplicar el MDL.

En este estudio sólo se tienen en cuenta los atributos con un peso superior a 0, descartando todos aquellos con valores 0 o negativos.

Finalmente, se han utilizado algoritmos de análisis de regresión para la generación del modelo, que es la predicción del objetivo, número de demandas sanitarias urgentes por enfermedades respiratorias que se solicitarán en una fecha determinada.

Una tarea de regresión comienza con un conjunto de datos en el que se conocen los valores objetivo. Un algoritmo de regresión estima el valor objetivo como una función de los predictores para cada instancia del conjunto de datos. Las relaciones entre los predictores y el objetivo se resumen en un modelo que puede aplicarse a otro conjunto de datos en el que se desconocen los valores objetivo. Los modelos de regresión se prueban calculando varios estadísticos que miden la diferencia entre los valores predichos y los reales [65,66].

Los datos históricos para un proyecto de regresión suelen dividirse en dos conjuntos de datos: uno para construir el modelo y otro para probar el modelo. Es necesario especificar

que la fecha utilizada para probar el modelo no se incluye en el conjunto de datos de entrenamiento.

Para este estudio, se utiliza un modelo lineal puro del algoritmo Generalized Linear Models (GLM) [79] y otros modelos que aplican Support Vector Machines (SVM) con Gaussian y Linear Kernel, respectivamente [67,80,81].

4. Resultados

El análisis previo del conjunto de datos de la sección 2.3 mostró que en este estudio es necesario separar la predicción de los días laborables de la de los días festivos o fines de semana.

La razón de ello es que los motivos que llevan a la demanda de servicios de emergencia son diferentes en ambos casos, por lo que se requiere un estudio separado.

4.1 Generación del modelo para días laborables.

Basándose en los algoritmos descritos anteriormente, se utiliza un conjunto de datos de 2012 a 2019 como datos de entrenamiento. A este respecto, es importante señalar que no se han incluido los datos correspondientes al periodo de la pandemia COVID-19. La razón de ello es que, durante esos años en Jaén, los servicios de atención primaria estaban cerrados y sólo se prestaba atención telefónica. Como resultado, los datos sobre demandas sanitarias urgentes para este periodo están anormalmente saturados. La inclusión de este periodo en el estudio distorsionaría la tendencia del modelo, ya que no existe homogeneidad en las condiciones en las que se produce la demanda de servicios de urgencias. Los modelos predictivos se diseñan utilizando los algoritmos mencionados en los apartados anteriores.

Hemos realizado ocho modelos diferentes para estos datos de jornada laboral, en cada uno de ellos hemos tomado un año

diferente para comprobar la bondad de ajuste del modelo diseñado.

En primer lugar, como punto de partida, se toman los años 2013 a 2019 como datos de entrenamiento para nuestro modelo, dejando 2012 como año de comparación para contrastar los resultados obtenidos en nuestro modelo con las demandas reales de recursos recibidas en 2012.

A continuación, para evaluar la fiabilidad del modelo, se utilizó la técnica de validación cruzada k-fold para evaluar los resultados en análisis estadísticos [82].

Esta metodología de evaluación consiste en valorar la calidad de la predicción de cada año separando los datos del año a predecir de los datos de entrenamiento.

En particular, para esta investigación, se utilizó para probar la fiabilidad del modelo en la predicción del número de demandas de emergencias sanitarias que se recibirían en el centro de vigilancia sanitaria.

Como se ha indicado anteriormente, para este estudio se utilizaron los datos de producción de 2012 a 2019.

La calidad teórica de los modelos generados a partir de cada uno de los algoritmos de regresión se evaluó a partir de las diferencias entre los valores predichos y los valores reales, como podemos ver en la tabla 19.

Algorithm	Mean Absolute Error	Root Mean Square Error	Mean Actual Value	Mean Predictive Value
GLM	1823	2207	6241	6649

SVM with lineal kernel	1.612	2025	6241	6326
SVM with Gaussian kernel	1629	2069	6241	6326

Tabla 19 Métricas de error de los modelos predictivos obtenidos para los días laborables utilizando 2012 como año de control.

La teoría nos dice que los resultados obtenidos por las dos variantes del algoritmo SVM son mejores que los obtenidos por el algoritmo GLM.

La tabla 20 muestra los parámetros utilizados para la aplicación del algoritmo SVM. Aunque los valores obtenidos en las dos variantes del algoritmo SVM son muy similares, se demostrará más adelante que si se utilizan los datos promediados sobre todos los años, el algoritmo SVM con kernel gaussiano dará un mejor resultado predictivo.

SVM with Gaussian kernel	Parameters
Standard Deviation	2028306
Complexity Factor	0.589921
Kernel Function	Gausiano
Algorithm Name	Support Vector Machine
Active Optimisation	Enable
Automatic Preparation	Enable
SVMS_EPSILON	0.040708
Core Cache Size	50,000.000
Tolerance	0.001

Tabla 20 Parámetros del algoritmo SVM con núcleo gaussiano utilizado para generar el modelo de predicción. Datos para la predicción de 2012.

En la configuración del modelo SVM, se ha especificado el parámetro del factor de complejidad.

Este parámetro permite equilibrar el error del modelo (medido con respecto a los datos de entrenamiento) y la complejidad del modelo para evitar un ajuste excesivo o insuficiente de los datos.

Los valores mayores penalizan más los errores, lo que conlleva un mayor riesgo de ajuste excesivo de los datos. Los valores menores penalizan menos los errores y pueden provocar un ajuste insuficiente.

Tras múltiples pruebas, este valor por defecto es 1 y en nuestro estudio se reduce a 0,589 precisamente para evitar el sobreajuste. Tabla 20.

La segunda parte de este análisis consiste en comprobar si los resultados teóricos obtenidos en la primera parte se corresponden con la realidad.

Como medida de la bondad de los modelos obtenidos, hemos tomado cada semana de 2012 y calculado el valor absoluto de la diferencia entre el número real de recursos demandados y el valor predicho por nuestros tres modelos.

Para confirmar que el algoritmo SVM con kernel gaussiano es el que mejor funciona para los días laborables, se calcula el valor medio de los errores absolutos de los tres modelos para todos los años, tabla 21.

2012-2019	Mean absolute error
-----------	---------------------

GLM	30.4%
SVM with lineal kernel	27.8%
SVM with Gaussian kernel	27.3%

Tabla 21 Error absoluto medio de los resultados para los modelos de días laborables.

Se confirma que el algoritmo SVM funciona mejor que el GLM. También se confirma que el algoritmo SVM con kernel gaussiano obtiene mejores resultados que el algoritmo lineal, aunque sin grandes diferencias.

Una vez comprobado que el algoritmo SVM con kernel gaussiano es el que mejor se ajusta, se analiza su comportamiento a lo largo de todas las semanas del año seleccionado, 2012, figura 19.

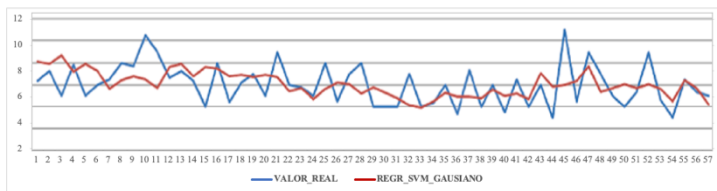


Figura 19 Comparación entre el valor real de las demandas de recursos sanitarios urgentes y el valor predicho por SVM con kernel gaussiano. Valores por semana considerando sólo los días laborables de 2012.

Los resultados mostrados en la Figura 19 confirman que la predicción del modelo SVM gaussiano es consistente en valores y tendencia a lo largo de todas las semanas de 2012.

En la primera parte del año el número de recursos de emergencia movilizados por enfermedades respiratorias fue superior al resto del año, debido a las bajas temperaturas y al aumento de la humedad en el ambiente.

A mediados de año se produjo un ligero descenso en el número de recursos necesarios en comparación con el comienzo del año.

Hay que tener en cuenta que, al tratarse de urgencias y emergencia, no cabe esperar un descenso muy pronunciado.

Por último, como era de esperar, en el último periodo del año el número de recursos volvió a aumentar ligeramente al empeorar de nuevo las condiciones meteorológicas, con la llegada del invierno.

Es interesante ver cómo la predicción de nuestro modelo se ajusta a la tendencia general a lo largo del año, así como los pequeños picos que se producen durante las semanas.

4.2 Generación de modelos en festivos y fines de semana

La metodología seguida es idéntica a la de los días laborables, con siete de los ocho años de datos de entrenamiento disponibles se utilizan para generar los modelos y se deja uno para comprobar la calidad de la predicción.

Al igual que en el caso de los días laborables, se han confeccionado ocho modelos, tomando en cada modelo un año de control distinto.

Para realizar una comparación coherente con el caso de los días laborables, se seleccionó de nuevo el año 2012 como año de control y el periodo comprendido entre 2013 a 2019 como datos de entrenamiento como modelo ejemplo para las explicaciones.

Los resultados teóricos sobre la calidad de los modelos obtenidos se muestran en la tabla 22.

En este caso, según las métricas de los modelos todos ellos obtienen resultados muy similares. Sin embargo, teóricamente, el que mejor se ajusta al valor real es el algoritmo SVM con kernel lineal con un error absoluto medio de 2,326.

Algorithm	Mean Absolute Error	Root Mean Square Error	Mean RealValue	Mean predictive Value
GLM	2.413	3.235	7.172	6.687
SVM with lineal kernel	2.326	3.213	7.172	6.185
SVM with Gaussian kernel	2.343	3.268	7.172	6.140

Tabla 22 Métricas de error de los modelos predictivos obtenidos para fines de semana y festivos utilizando 2012 como año de control.

Para corroborar si el algoritmo SVM con kernel lineal es el que obtiene los mejores resultados para fines de semana y festivos, se calcula el valor medio de los errores absolutos de los tres

algoritmos para todos los años utilizados como conjunto de datos de entrenamiento.

Así, se puede confirmar que en esta ocasión los modelos se comportan de forma muy similar, aunque como puede observarse en la tabla 23, el valor medio correspondiente al modelo SVM con kernel lineal es ligeramente mejor que el de los otros dos.

2012-2019	Mean absolute error
GLM	34.3%
SVM with lineal kernel	32.8%
SVM with Gaussian kernel	33.0%

Tabla 23 Error absoluto medio de los resultados de los modelos de fines de semana y festivos.

La figura 20 permite un análisis gráfico detallado del comportamiento predictivo del modelo generado con el algoritmo SVM con kernel lineal, que, como se ha demostrado, ofrece los mejores resultados.

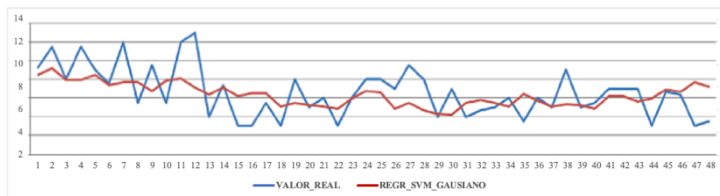


Figura 20 Comparación entre el valor real de las demandas de recursos sanitarios urgentes y el valor predicho por SVM con kernel gaussiano. Valores por semana considerando fines de semana y festivos en 2012.

Este algoritmo se ajusta a la tendencia de los recursos solicitados a lo largo de 2012, identificando incluso pequeños picos a lo largo de este año.

Al igual que ocurría los días laborables, la tendencia de los recursos necesarios los fines de semana disminuye ligeramente a medida que avanza el año y mejoran las condiciones meteorológicas.

Cabe mencionar la última semana del año, en la que la diferencia entre los valores reales y la previsión es ligeramente diferente. Si se analiza en detalle, se debe a que en esta última semana hubo un número extraordinariamente reducido de solicitudes de ayuda de emergencia en comparación con el mismo periodo de cualquier otro año.

Otro aspecto positivo es que la predicción de nuestro modelo detecta el pequeño cambio en la tendencia temporal de los casos reales a mediados de año.

Otra evaluación se realiza analizando los coeficientes resultantes para la regresión lineal del modelo predictivo SVM con kernel lineal, tabla 24.

Variable	Value	Coefficients
maximum_humidity		-0.028
mean_humidity		0.093
minimum_humidity		0
month	10	-0.086
month	1	0.079
month	11	-0.059
month	5	-0.049
month	9	0.038
month	3	0.034
month	4	0.029
month	8	0.029
month	12	-0.024
month	2	0.019
month	7	-0.01
n_days	1	-0.041
n_days	4	0.029
n_days	2	0.021
n_days	3	-0.009
precipitation		0.019
radiation		0.09
week_month	2	0.019
week_month	4	-0.017
week_month	5	0.008
week_month	3	-0.007
week_month	1	-0.003
maximum_temperature		-0.22
mean_temperature		0.039
minimum_temperature		-0.006
wind_speed		0.072

Tabla 24 Coeficientes utilizados para el algoritmo SVM Lineal

Estos coeficientes indican el nivel de influencia de cada variable sobre el atributo objetivo.

Cabe señalar que este modelo toma como datos de aprendizaje los correspondientes a fines de semana y festivos, un volumen mucho menor que el modelo de días laborables, pero necesario debido al diferente funcionamiento de la demanda de recursos.

Si se analiza la variable Mes, se observa que los tres o cuatro primeros meses del año tienen un mayor peso que los meses siguientes, esto es debido a que son meses con condiciones climatológicas adversas, lo que aumenta la presencia de enfermedades respiratorias, dando lugar a un mayor número de datos disponibles para el modelo.

También influye en la mayor disponibilidad de información el mayor número de días festivos en la ciudad de Jaén en los cuatro primeros meses del año.

También se observa que las semanas con 4 días festivos tienen un mayor peso en el modelo que las semanas con menos días, ya que disponemos de más información de entrenamiento coincidiendo con los periodos vacacionales.

Por último, se observa que las variables humedad media, temperatura media y velocidad del viento tienen un peso positivo muy importante. Esto se debe a que tienen una influencia directa en los síntomas relacionados con enfermedades respiratorias, como resfriados, bronquitis, asma provocada por el polen alergias, etc...

Aunque estos datos corresponden al modelo en el que el año de control es 2012, se ha verificado un comportamiento similar en los modelos correspondientes a los otros siete años restantes.

5. Diseño del modelo de datos

La creación de un modelo consiste en elegir una forma de diseñar nuestros esquemas de base de datos. Posteriormente tendremos que elegir un gestor de base de datos determinado en el que darle forma a dicho proyecto.

En nuestra investigación para la creación de base de datos hemos elegido SGBD Oracle.

Con Oracle SQL Developer hemos creado toda la estructura de tablas que darán soporte a los futuros modelos predictivos. También hemos creado y ejecutado todas las consultas que han ido dando forma a las tablas definitivas utilizadas en nuestros modelos.

Con Oracle Data Mining hemos realizado toda la investigación, creación, ejecución y depuración de todos nuestros modelos. Hemos creado modelos para cada uno de los algoritmos estudiados y citados con anterioridad.

Después de la investigación realizada sobre cada uno de los algoritmos que pueden ser de interés para nuestro proyecto, pasamos a la creación de los mismos. Como se ha comentado con anterioridad dichos modelos se han creado con Oracle Data Mining mediante la herramienta SQL Developer.

En primer lugar construimos la tabla definitiva sobre la que vamos a ejecutar nuestro modelo. Dicha tabla contiene representación de todas las áreas de conocimiento que queremos integrar. Dicha información ha sufrido diversos procesos de depuración, agrupamiento y preparación para que

el resultado final sea el más apropiado para la aplicación de los modelos a desarrollar.

Tras un largo proceso de creación y desarrollo de las tablas, se han creado 16 modelos predictivos, la mitad aplicando el filtro de los días festivos y fines de semana y la otra mitad para los días laborables.

El tipo de modelo creado ha sido un modelo de regresión. En la construcción de dicho modelo seleccionamos el atributo destino o target sobre el que necesitamos la predicción y todos aquellos atributos que utilizará dicho modelo predictivo para obtener el mejor resultado final.

Los algoritmos utilizados en la construcción de nuestros modelos han sido los referenciados con anterioridad, GLM y SVM con nucleo lineal y gaussiano.

Tal y como hemos comentado en el punto anterior, para los datos de los días festivos el modelo que mejor resultados nos ofrece es el que utiliza el algoritmo SVM lineal, aunque ofrece resultados muy parecidos al SVM gaussiano. En ambos casos el algoritmo SVM funciona mejor que el algoritmo GLM.

En el caso de los modelos para los días laborables volvemos a confirmar que el algoritmo SVM es mejor que el GLM, pero en este caso el algoritmo SVM gaussiano ofrece mejores resultados que el lineal.

A continuación, en la figura 21 ver una imagen ejemplo de uno de los modelos.

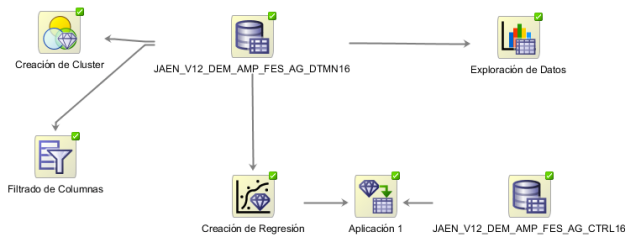


Figura 21 Ejemplo de modelo predictivo de regresión.

En este modelo podemos ver como por un lado tenemos la tabla `JAEN_V12_DEM_AMP_FES_AG_DTMN16`, que contiene todos los datos históricos, de todas las áreas utilizadas, debidamente enlazados y organizados. De todos estos datos, se ha eliminado un año de información que será utilizado con posterioridad para contrastar la eficacia de nuestro modelo.

Esta tabla referenciada anteriormente, constituye la fuente de datos de entrada de nuestro sistema de regresión. A los cuales se les aplicará los correspondientes algoritmos utilizados por nuestro sistema.

Por otro lado tenemos la tabla `JAEN_V12_DEM_AMP_FES_AG_CTRL16`, dicha tabla tiene el año de datos históricos reales que eliminamos de la tabla anterior. De esta forma, la predicción de nuestro sistema para dicho año, es comparada con el valor real que se produjo en ese mismo periodo. De esta forma podemos comprobar el grado de eficacia de nuestro sistema de regresión, ya que podemos ver de que forma se ha ajustado la predicción al valor real.

Por último, tenemos que tener en cuenta que para un aprendizaje continuo de nuestro modelo iremos alimentando la tablas utilizadas con los nuevos datos reales que se vayan produciendo. De esta forma nuestro modelo tendrá un aprendizaje continuo y mejoraremos los resultados de forma continua en el tiempo.

6. Discusión

En nuestro proyecto hemos querido abordar un tema de vital importancia para nosotros como es la aparición y gestión de urgencias y emergencias sanitarias en la ciudad de Jaén.

Al ser un tema tan sensible para todos los grupos de edad de la población, hemos estudiado como nos podríamos anticipar, utilizando minería de datos y algoritmos predictivos, a la aparición de dichas asistencias para así poder mejorar en la medida de lo posible la gestión de la resolución de las mismas.

Hemos estudiado la dinámica de aparición de urgencias y emergencias en la población de Jaén, teniendo en cuenta que al ser situaciones sobrevenidas dicho estudio no parece trivial.

Para obtener conclusiones concretas, realizamos el estudio sobre los recursos movilizados para urgencias y emergencias de patologías respiratorias en la población de la ciudad de Jaén.

Observamos con dicho estudio que la aparición de este tipo de asistencias es distinto los fines de semana y festivos que los días laborales. Esto puede ser debido a que durante la semana hay pacientes que hacen uso de los servicios de atención primaria en un primer momento, evitando así un posible empeoramiento de su salud, no llegando a convertirse en una verdadera urgencia o emergencia.

También puede ser debido a que entre semana hay un tipo de urgencias ambulatorias que no todas existen los fines de semana.

Otro aspecto importante que nos despertó interés es confirmar qué aspectos afectan en mayor o menor medida a la aparición de dichas asistencias.

Mediante el algoritmo MDL hemos obtenido el grado de importancia o decisión de cada uno de los atributos utilizados en la resolución del objetivo.

Ya en la aplicación de los algoritmos predictivos GLM y SVM, hemos podido confirmar que los atributos velocidad del viento, temperatura y humedad media son de gran importancia en la aparición de asistencias respiratorias. Se ha podido confirmar como los meses con peores condiciones meteorológicas y alérgicos son importantes para este tipo de problemas respiratorios, siendo prácticamente nulos para nuestro modelo meses como agosto.

Hemos comprobado cómo el mayor número de datos de días festivos en la primera mitad del año y en el último mes de diciembre mejoran positivamente la predicción de nuestro sistema, al tener más datos de entrenamiento.

Como era de esperar la aparición de casos reales es más alta los primeros 4 meses del año, bajando el número de incidencias en los meses centrales y subiendo ya prácticamente al final del mismo.

Hemos comprobado como nuestro sistema se ajusta perfectamente a esta dinámica, pero nos ha resultado muy interesante comprobar como nuestro modelo detecta los pequeños picos puntuales que se producen en los meses centrales, prediciendo pequeñas subidas en esas mismas semanas.

Al tener la ciudad de Jaén dividida en cuadrantes geolocalizados de 250mx250m, podríamos dotar a la ciudad de Jaén de las unidades de emergencias necesarias en cada época del año y ubicarlas en las zonas con mayor previsión de aparición de asistencias.

Con todo esto conseguiríamos un ahorro económico al ajustar los equipos tanto humanos como materiales a las necesidades reales y lo que es más importante al ubicarlos en las zonas donde existe mayor previsión de aparición de asistencias estamos minimizando el tiempo de espera del usuario a ser atendido.

Al ser una patología tiempo-dependiente es un aspecto de vital importancia, consiguiendo así un beneficio directo en la salud de la ciudadanía.

7. Conclusiones

En este estudio se ha conseguido el reto de generar un modelo predictivo eficaz para estimar las demandas sanitarias de urgencias y emergencias que se recibirán en el centro de control del municipio de Jaén, España.

En concreto en nuestro caso se ha calculado para la casuística específica más recurrente en este municipio, que son las urgencias y emergencias relacionadas con enfermedades respiratorias.

Además, se han generado dos modelos de predicción, uno para días laborables y otro para fines de semana y festivos, ya que el comportamiento de ambos modelos es diferente.

Los resultados indican que el mejor algoritmo para predecir los recursos sanitarios necesarios para atender los episodios de urgencias respiratorias en la ciudad de Jaén en días laborables es el algoritmo SVM con kernel Gaussiano. Sin embargo, para el caso de fines de semana y festivos, el algoritmo SVM con kernel lineal resultó más eficiente.

Los coeficientes de regresión de los modelos muestran una relación directa entre las condiciones meteorológicas y el aumento de pacientes con insuficiencia respiratoria y, en consecuencia, con la necesidad de incrementar los recursos de emergencia para atender esta demanda.

En la tabla 25 de abajo, se observa que las variables humedad media, temperatura media y velocidad del viento tienen un peso positivo importante. Esto se debe a que tienen una influencia directa en los síntomas relacionados con

enfermedades respiratorias, como resfriados, bronquitis, asma, etc... Figura 22.

Datos atmosféricos	
Temperatura Media	0,039
Humedad media	0,093
Velocidad viento	0,072

Tabla 25 Principales variables con peso positivo

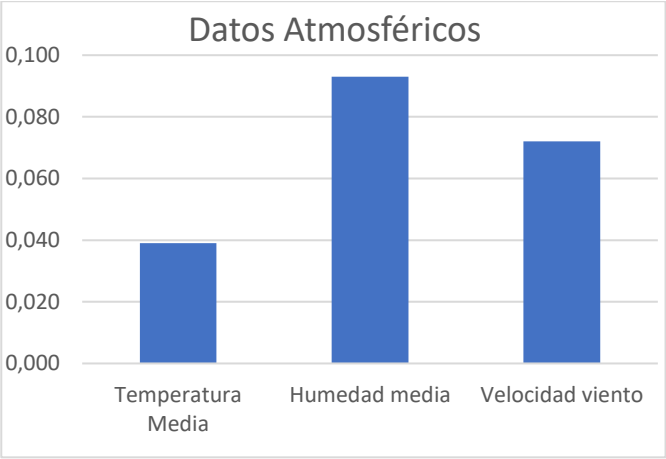


Figura 22 Principales variables con peso positivo.

Como podemos ver en la tabla 26, en la aplicación del algoritmo SVM Lineal en los fines de semana y festivos, en los primeros meses del año existe una mayor tendencia a la aparición de patologías respiratorias, ya que las malas

condiciones atmosféricas de estos meses favorecen la aparición de estas patologías.

Podemos ver como en los meses centrales del año este atributo pierde importancia y a finales de año vuelve a ascender de nuevo al ir empeorando las condiciones atmosféricas, figura 23.

Mes	Coeficiente SVM Lineal
Enero	0,079
Febrero	0,019
Marzo	0,034
Abril	0,029
Mayo	-0,049
Junio	-0,020
Julio	-0,010
Agosto	0,029
Septiembre	0,038
Octubre	-0,086
Noviembre	-0,059
Diciembre	-0,024

Tabla 26 Coeficientes de la variable mes con SVM Kernel Lineal

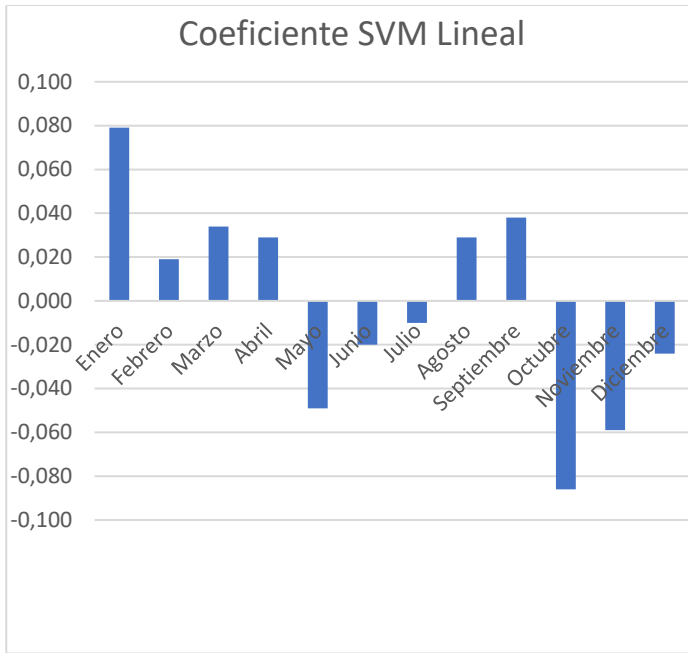


Figura 23 Coeficientes de la variable mes con SVM Kernel Lineal.

El principal logro de esta investigación es haber construido un modelo capaz de predecir el número de recursos sanitarios de emergencia necesarios para atender a pacientes con enfermedades respiratorias en una zona y periodo determinados.

Al disponer de la información georeferenciada, además de predecir número de recursos necesarios, podemos saber que zona de la ciudad de Jaén necesita fortalecer su flota de recursos y que otra zona puede estar sobredimensionada.

Por tanto, este modelo ayuda a tomar decisiones sobre el dimensionamiento eficiente de los recursos sanitarios disponibles y su localización más optima dentro de la ciudad de Jaén.

El uso de este modelo evita la escasez y el sobredimensionamiento de los recursos sanitarios humanos y materiales.

Por otro lado, permite ubicar los recursos donde realmente van a ser necesarios para minimizar el tiempo de espera de los pacientes, aspecto vital en las emergencias, que influye directa y positivamente en su salud.

Por otro lado, desde un punto de vista menos vital pero también importante, la utilización de este modelo predictivo tiene un impacto positivo en los costes económicos derivados de la gestión de las urgencias, ya que permite ajustar los recursos sanitarios humanos y materiales a los estrictamente necesarios.

En consecuencia, el modelo predictivo generado es una herramienta vital para gestionar y dar respuesta en el menor tiempo posible a los pacientes con enfermedades respiratorias que requieren atención urgente en la ciudad de Jaén.

8. Posibles nuevos estudios

Con este trabajo hemos comprobado la eficacia de los modelos predictivos para mejorar la calidad del paciente en cuanto a la gestión de recursos humanos y técnicos en urgencias y emergencias.

8.1 Ampliación a otras demandas asistenciales

Las demandas realizadas por los usuarios a los teléfonos de urgencias y emergencias se clasifican en distintos tipos, informativas, asistenciales, transporte, etc...

En concreto en este estudio nos hemos centrado en las demandas asistenciales y dentro de estas hemos analizado las patologías respiratorias.

Una futura línea de trabajo muy interesante será realizar nuevos modelos para analizar otros grupos de patologías importantes, como puede ser problemas en pacientes cardíacos o con pérdida de consciencia.

8.2 Red de Transporte Urgente

Una nueva línea de trabajo sería crear modelos predictivos para optimizar la red de transporte urgente de la provincia de Jaén. La red de transporte urgente de la provincia de Jaén consta de cerca 50 ambulancias. Dichas ambulancias son gestionadas

para dar respuesta al traslado urgente de pacientes desde el lugar de la asistencia hasta los centros hospitalarios.

Estos traslados se realizan en aquellos casos en los que no es necesario o no existe un recurso con un equipo médico completo.

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

9. Difusión del estudio

Este trabajo se ha publicado en los siguientes medios.

Congreso_____

CISTI 2022 - 17ª Conferencia Ibérica de Sistemas y Tecnologías de Información

Fecha y Lugar: 22-25 Junio 2022. Madrid

Título: Optimización de La Calidad Asistencial de Urgencias y Emergencias Mediante Bases de Datos Multidisciplinares

Detalle:

J. C. Arias, J. J. Cubillas, M. I. Ramos and F. R. Feito,

"Optimization of the quality of emergency care through multidisciplinary databases," 2022 17th Iberian Conference on Information Systems and Technologies (CISTI), Madrid, Spain, 2022, pp. 1-5,

doi: 10.23919/CISTI54924.2022.9820109.

Abstract:

At present, one of the most important techniques for improving the quality of patient care is the exhaustive analysis of all aspects of such care through the use of specific and detailed databases. A database capable of integrating and relating all the aspects that directly and indirectly affect the assistance or emergencies carried out in the province of Jaén (Jaen city, Andalusian, Spain), over a period of eight years, has been created. For this purpose, health, social, economic, environmental and geographical information has been stored

and related to each assistance. The high level of interrelation that exists between all the variables involved in patient care has been used to analyse and predict the number and type of care that will appear in the future, as well as to efficiently manage all the resources involved in such care.

keywords: Economics, Databases, Urban areas, Optimization, Information systems, Healthcare, Database design, Geospatial data,

URL:

<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9820109&isnumber=9819977>

Calidad:

Indexado en Scopus y Web of Science (WoS)

Base de datos de indexación: JCR (Journal Citation Reports, SCIE)

Categorías: Computer Science, Information Systems.

Artículo_____

Revista: Electronics

Título: Optimising Health Emergency Resource Management from Multi-Model Databases

Detalle:

Arias, Juan C., Juan J. Cubillas, and Maria I. Ramos. "Optimising Health Emergency Resource Management from Multi-Model Databases" Electronics 11, 2022. no. 21: 3602.

<https://doi.org/10.3390/electronics11213602>

Abstract:

The health care sector is one of the most sensitive sectors in our society, and it is believed that the application of specific and detailed database creation and design techniques can improve the quality of patient care. In this sense, better management of emergency resources should be achieved. The development of a methodology to manage and integrate a set of data from multiple sources into a centralised database, which ensures a high quality emergency health service, is a challenge. The high level of interrelation between all of the variables related to patient care will allow one to analyse and make the right strategic decisions about the type of care that will be needed in the future, efficiently managing the resources involved in such care. An optimised database was designed that integrated and related all aspects that directly and indirectly affected the emergency care provided in the province of Jaén (city of Jaén, Andalusia, Spain) over the last eight years. Health, social, economic, environmental, and geographical information related to each of these emergency services was stored and related. Linear and nonlinear regression algorithms were used: support vector machine (SVM) with linear kernel and generated linear model (GLM), and the nonlinear SVM with Gaussian kernel. Predictive models of emergency demand were generated with a success rate of over 90%.

Keywords: healthcare; database design; geospatial data

<https://www.mdpi.com/2079-9292/11/21/3602>.

Calidad:

Indexado en Scopus y Web of Science (WoS)

Base de datos de indexación: JCR (Journal Citation Reports, SCIE)

Índice de Impacto: 2.9 (Año 2022)

Categorías:

Computer Science Q3, Information SystemsEngineering,
Electrical & Electronic Q2 and Physics Applied Q2.

Citas: 2

Artículo_____

Revista: International Journal of Medical Informatics

Título: Predicting emergency health care demands due to
respiratory diseases

Detalle:

J.C. Arias, M.I. Ramos, J.J. Cubillas,

“Predicting emergency health care demands due to respiratory
diseases”

International Journal of Medical Informatics, Volume 177,
2023, 105163, ISSN 1386-5056,

<https://doi.org/10.1016/j.ijmedinf.2023.105163>.

Abstract:

Background

Timely care in the health sector is essential for the recovery of patients, and even more so in the case of a health emergency. In these cases, appropriate management of human and technical resources is essential. These are limited and must be mobilised in an optimal and efficient manner.

Objective

This paper analyses the use of the health emergency service in a city, Jaén, in the south of Spain. The study is focused on the most recurrent case in this service, respiratory diseases.

Methods

Machine Learning algorithms are used in which the input variables are multisource data and the target attribute is the prediction of the number of health emergency demands that will occur for a selected date. Health, social, economic, environmental, and geospatial data related to each of the emergency demands were integrated and related. Linear and nonlinear regression algorithms were used: support vector machine (SVM) with linear kernel and generated linear model (GLM), and the nonlinear SVM with Gaussian kernel.

Results

Predictive models of emergency demand due to respiratory diseases were generated with an absolute error better than 35 %.

Conclusions

This model helps to make decisions on the efficient sizing of emergency health resources to manage and respond in the shortest possible time to patients with respiratory diseases requiring urgent care in the city of Jaén.

Keywords: Machine Learning; Prediction; Health Emergency Service; Geospatial data

(<https://www.sciencedirect.com/science/article/pii/S1386505623001818>)

Calidad:

Indexado en Scopus y Web of Science (WoS)

Base de datos de indexación: JCR (Journal Citation Reports, SCIE)

Índice de Impacto: 4.9 (Año 2022)

Categorías:

Computer Science, Information Systems Q2, Health Care Sciences & Services Q1 and Medical Informatics Q2.

Citas: 1

Poster_____

Evento: Jornadas Doctorales Universidad de Jaén

Título: Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

Fecha y Lugar: Junio 2022, Baeza (Jaén).

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

10. Bibliografía

1. *Institute of Medicine (US). Improving the Nation's Health Care System; National Academies Press (US): Washington, DC, USA, 2009.*
2. Nundy, S., Cooper, L. A., & Mate, K. S. (2022). The quintuple aim for health care improvement: a new imperative to advance health equity. *Jama*, 327(6), 521-522.
3. Vaz, N.; Venkatesh, R. *Service Design in the Healthcare Space with a Special Focus on Non-Clinical Service Departments: A Synthesis and Future Directions*. *Health Serv. Manag. Res.* 2022, 35, 83–91.
4. Ramos, I., Cubillas, J. J., Feito, F. R., & Ureña, T. (2016). Spatial analysis and prediction of the flow of patients to public health centres in a middle-sized Spanish city. *Geospatial Health*, 11(3).
5. Ganguly, A.; Nandi, S. *Using Statistical Forecasting to Optimize Staff Scheduling in Healthcare Organizations*. *J. Health Manag.* 2016, 18, 172–181.
6. Youn, S., Geismar, H. N., & Pinedo, M. (2022). Planning and scheduling in healthcare for better care coordination: Current understanding, trending topics, and future opportunities. *Production and Operations Management*, 31(12), 4407-4423.
7. Wiréhn, A.-B.E.; Karlsson, H.M.; Carstensen, J.M. *Estimating Disease Prevalence Using a Population-Based Administrative Healthcare Database*. *Scand. J. Public Health* 2007, 35, 424–431.

8. Hittle, M., Culpepper, W. J., Langer-Gould, A., Marrie, R. A., Cutter, G. R., Kaye, W. E., ... & Wallin, M. T. (2023). *Population-Based Estimates for the Prevalence of Multiple Sclerosis in the United States by Race, Ethnicity, Age, Sex, and Geographic Region*. *JAMA neurology*.
9. Kerr, K.; Norris, T.; Stockdale, R. *Data Quality Information and Decision Making: A Healthcare Case Study*. In *Proceedings of the 18th Australasian Conference on Information Systems, Toowoomba, Australia, 5–7 December 2007*.
10. Zozus, M. N., Kahn, M. G., & Weiskopf, N. G. (2023). *Data quality in clinical research*. In *Clinical research informatics* (pp. 169-198). Cham: Springer International Publishing.
11. Salman, O.H.; Rasid, M.F.A.; Saripan, M.I.; Subramaniam, S.K. *Multi-Sources Data Fusion Framework for Remote TriagePrioritization in Telehealth*. *J. Med. Syst.* 2014, 38, 103.
12. Salman, O. S., Abdul Latiff, N. M. A., Salman, O. H., & Syed Ariffin, S. H. (2024). *A hybrid computational approach to process real-time streaming multi-sources data and improve classification for emergency patients triage services: moving forward to an efficient IoMT-based real-time telemedicine systems*. *Neural Computing and Applications*, 1-14.
13. Trifirò, G.; Coloma, P.M.; Rijnbeek, P.R.; Romio, S.; Mosseveld, B.; Weibel, D.; Bonhoeffer, J.; Schuemie, M.; van der Lei, J.;Sturkenboom, M. *Combining Multiple Healthcare Databases for Postmarketing Drug and*

Vaccine Safety Surveillance: Why and How? J. Intern. Med. 2014, 275, 551–561.

14. Ramos, M.I.; Cubillas, J.J.; Feito, F.R. Improvement of the Prediction of Drugs Demand Using Spatial Data Mining Tools. J. Med. Syst. 2016, 40, 6.
15. Burkle, F.M.; Bradt, D.A.; Ryan, B.J. Global Public Health Database Support to Population-Based Management of Pandemics and Global Public Health Crises, Part I: The Concept. Prehospital Disaster Med. 2021, 36, 95–104.
16. Khorram-Manesh, A., Goniewicz, K., & Burkle Jr, F. M. (2023). Unleashing the global potential of public health: a framework for future pandemic response. Journal of Infection and Public Health.
17. Mezghani, E.; Exposito, E.; Drira, K.; Da Silveira, M.; Pruski, C. A Semantic Big Data Platform for Integrating Heterogeneous Wearable Data in Healthcare. J. Med. Syst. 2015, 39, 185.
18. Savoska, S., Ristevski, B., Blazheska-Tabakovska, N., Jolevski, I., Bocevska, A., & Trajkovik, V. (2023). Integration of heterogeneous medical and biological data with electronic personal health records. Medical Informatics, Biometry and Epidemiology (MIBE), 19, 1-11.
19. Wang, Y.; Kung, L.; Byrd, T.A. Big Data Analytics: Understanding Its Capabilities and Potential Benefits for Healthcare Organizations. Technol. Forecast. Soc. Chang. 2018, 126, 3–13.
20. Fanelli, S., Pratici, L., Salvatore, F. P., Donelli, C. C., & Zangrandi, A. (2023). Big data analysis for decision-

making processes: challenges and opportunities for the management of health-care organizations. Management Research Review, 46(3), 369-389.

21. Hamami, D.; Atmani, B.; Cameron, R.; Pollock, K.G.; Shankland, C. *Improving Process Algebra Model Structure and Parameters in Infectious Disease Epidemiology through Data Mining. J. Intell. Inf. Syst.* 2019, 52, 477–499.
22. Benhar, H.; Idri, A.; Fernández-Alemán, J.L. *A Systematic Mapping Study of Data Preparation in Heart Disease Knowledge Discovery. J. Med. Syst.* 2018, 43, 17.
23. Werner de Vargas, V., Schneider Aranda, J. A., dos Santos Costa, R., da Silva Pereira, P. R., & Victória Barbosa, J. L. (2023). *Imbalanced data preprocessing techniques for machine learning: a systematic mapping study. Knowledge and Information Systems*, 65(1), 31-57.
24. Uthiramoorthy, A., Bhardwaj, A., Singh, J., Pant, K., Tiwari, M., & Gonzáles, J. L. A. (2023, January). *A Comprehensive review on Data Mining Techniques in managing the Medical Data cloud and its security constraints with the maintained of the communication networks. In 2023 International Conference on Artificial Intelligence and Smart Communication (AISC) (pp. 618-623). IEEE.*
25. Oussous, A.; Benjelloun, F.-Z.; Ait Lahcen, A.; Belfkih, S. *Big Data Technologies: A Survey. J. King Saud Univ. Comput. Inf. Sci.* 2018, 30, 431–448.
26. Lee, C.H.; Yoon, H.-J. *Medical Big Data: Promise and Challenges. Kidney Res. Clin. Pract.* 2017, 36, 3–11.

27. *UNDP. Human Development Report 2015; UNDP: New York, NY, USA, 2015.*
28. *Alkhwaldi, A. F. (2024). Understanding the acceptance of business intelligence from healthcare professionals' perspective: An empirical study of healthcare organizations. International Journal of Organizational Analysis.*
29. *Guo, C., & Chen, J. (2023). Big data analytics in healthcare. In Knowledge Technology and Systems: Toward Establishing Knowledge Systems Science (pp. 27-70). Singapore: Springer Nature Singapore..*
30. *Getz, K., Hubbard, R. A., & Linn, K. A. (2023). Performance of multiple imputation using modern machine learning methods in electronic health records data. Epidemiology, 34(2), 206-215.*
31. *Yue, X.; Wang, H.; Jin, D.; Li, M.; Jiang, W. Healthcare Data Gateways: Found Healthcare Intelligence on Blockchain with Novel Privacy Risk Control. J. Med. Syst. 2016, 40, 218.*
32. *Verma, G. (2024). Blockchain-based privacy preservation framework for healthcare data in cloud environment. Journal of Experimental & Theoretical Artificial Intelligence, 36(1), 147-160.*
33. *Goli-Malekabadi, Z.; Sargolzaei-Javan, M.; Akbari, M.K. An Effective Model for Store and Retrieve Big Health Data in Cloud Computing. Comput. Methods Programs Biomed. 2016, 132, 75–82.*
34. *Thatikonda, R., Padthe, A., Vaddadi, S. A., & Arnepalli, P. R. R. (2023). Effective Secure Data Agreement*

Approach-based cloud storage for a healthcare organization. International Journal of Smart Sensor and Adhoc Network, 60-70.

35. Graham, B.; Bond, R.; Quinn, M.; Mulvenna, M. *Using Data Mining to Predict Hospital Admissions From the Emergency Department. IEEE Access 2018, 6, 10458–10469.*
36. Kishore, K., Braitberg, G., Holmes, N. E., & Bellomo, R. (2023). *Early prediction of hospital admission of emergency department patients. Emergency Medicine Australasia.*
37. Gurazada, S.G.; Gao, S. (Caddie); Burstein, F.; Buntine, P. *Predicting Patient Length of Stay in Australian Emergency Departments Using Data Mining. Sensors 2022, 22, 4968.*
38. Wretborn, J., Wilhelms, D. B., & Ekelund, U. (2023). *Emergency department crowding and mortality: an observational multicenter study in Sweden. Frontiers in Public Health, 11, 1198188.*
39. Fekonja, Z., Kmetec, S., Fekonja, U., Mlinar Reljić, N., Pajnikihar, M., & Strnad, M. (2023). *Factors contributing to patient safety during triage process in the emergency department: A systematic review. Journal of Clinical Nursing.*
40. Beltrán Guzmán, I., Gil Cuesta, J., Trelles, M., Jaweed, O., Cherestal, S., van Loenhout, J. A. F., & Guha-Sapir, D. (2019). *Delays in arrival and treatment in emergency departments: Women, children and non-trauma consultations the most at risk in humanitarian settings. PLoS One, 14(3), e0213362.*

41. *Habebh, H., & Gohel, S. (2021). Machine learning in healthcare. Current Genomics, 22(4), 291.*
42. *Ramos, M. I., Cubillas, J. J., Jurado, J. M., Lopez, W., Feito, F. R., Quero, M., & Gonzalez, J. M. (2019). Prediction of the increase in health services demand based on the analysis of reasons of calls received by a customer relationship management. The International Journal of Health Planning and Management, 34(2), e1215-e1222.*
43. *Cubillas, J. J., Ramos, M. I., & Feito, F. R. (2022). Use of Data Mining to Predict the Influx of Patients to Primary Healthcare Centres and Construction of an Expert System. Applied Sciences, 12(22), 11453.*
44. *Gao, S., & Wang, H. (2022). Scenario prediction of public health emergencies using infectious disease dynamics model and dynamic Bayes. Future Generation Computer Systems, 127, 334-346.*
45. *Cubillas, J. J., Ramos, M. I., Gutiérrez, M. C., Feito, F. R., Parra, A., & Arias, J. C. (2015). Use of meteorological, environmental and spatial variables to predict drug use. In Digital Healthcare Empowering Europeans (pp. 938-938). IOS Press.*
46. *Qiu, H., Tan, K., Long, F., Wang, L., Yu, H., Deng, R., ... & Pan, J. (2018). The burden of COPD morbidity attributable to the interaction between ambient air pollution and temperature in Chengdu, China. International journal of environmental research and public health, 15(3), 492.*
47. *Duan, R. R., Hao, K., & Yang, T. (2020). Air pollution and chronic obstructive pulmonary disease. Chronic diseases and translational medicine, 6(4), 260-269.*

48. Liu, T., Jiang, Y., Hu, J., Li, Z., Li, X., Xiao, J., ... & Wang, Y. (2023). Joint associations of short-term exposure to ambient air pollutants with hospital admission of ischemic stroke. *Epidemiology*, 34(2), 282-292.
49. Humphrey-Murto, S., Makus, D., Moore, S., Watanabe Duffy, K., Maniate, J., Scowcroft, K., ... & Rangel, J. C. (2023). Training physicians and residents for the use of Electronic Health Records—A comparative case study between two hospitals. *Medical Education*, 57(4), 337-348.
50. Siddiqui, M. K., Morales-Menendez, R., Huang, X., & Hussain, N. (2020). A review of epileptic seizure detection using machine learning classifiers. *Brain informatics*, 7(1), 1-18.
51. Woldaregay, A. Z., Årsand, E., Botsis, T., Albers, D., Mamykina, L., & Hartvigsen, G. (2019). Data-driven blood glucose pattern classification and anomalies detection: machine-learning applications in type 1 diabetes. *Journal of medical Internet research*, 21(5), e11030.
52. Tian, L., Zhang, D., Bao, S., Nie, P., Hao, D., Liu, Y., ... & Wang, H. (2021). Radiomics-based machine-learning method for prediction of distant metastasis from soft-tissue sarcomas. *Clinical Radiology*, 76(2), 158-e19.
53. Kaouk, J. H., Garisto, J., Eltemamy, M., & Bertolo, R. (2019). Robot-assisted surgery for benign distal ureteral strictures: step-by-step technique using the SP® surgical system. *BJU international*, 123(4), 733-739.

54. Cubillas, J. J., Ramos, M. I., Feito, F. R., & Ureña, T. (2014). *An improvement in the appointment scheduling in primary health care centers using data mining*. *Journal of medical systems*, 38, 1-10.
55. Centro Emergencias Sanitarias. 061 | *Gestión de las Emergencias y Urgencias Sanitarias en Andalucía*; CES, 2024.
56. Arias, J. C., Cubillas, J. J., & Ramos, M. I. (2022). *Optimising Health Emergency Resource Management from Multi-Model Databases*. *Electronics*, 11(21), 3602.
57. REDIAM Red de Información Ambiental de Andalucía - Portal Ambiental de Andalucía Available online: <https://www.juntadeandalucia.es/medioambiente/portal/acceso-rediam>.
58. Sede Electrónica Del Catastro - Inicio Available online: <http://www.sedecatastro.gob.es/>.
59. INE INE. Instituto Nacional de Estadística Available online: <https://www.ine.es/>.
60. Instituto de Estadística y Cartografía de Andalucía Available online: <https://www.juntadeandalucia.es/institutodeestadisticaycartografia>.
61. Peña, D. (2023). *Detecting outliers and influential and sensitive observations in linear regression*. *Springer Handbook of Engineering Statistics*, 605-619.
62. Verma, D., Berwal, P., Khan, M. A., Alharbi, R. S., Alfaisal, F. M., & Rathnayake, U. (2023). *Design for the Prediction of Peak Outflow of Embankment Breaching*

Due to Overtopping by Regression Technique and Modelling. Water, 15(6), 1224..

63. Tang, W., He, H., & Tu, X. M. (2023). *Applied categorical and count data analysis*. CRC Press.
64. Dobson, A. J., & Barnett, A. G. (2018). *An introduction to generalized linear models*. CRC press.
65. Nakagawa, S., Yang, Y., Macartney, E. L., Spake, R., & Lagisz, M. (2023). Quantitative evidence synthesis: a practical guide on meta-analysis, meta-regression, and publication bias tests for environmental sciences. *Environmental Evidence*, 12(1), 1-19.
66. Taimal Yepes, C. A. (2023). Genetic algorithm with a Bayesian approach for multiple change-point detection in time series of counting exceedances for specific thresholds.
67. Yang, W., Li, C., & Jiang, L. (2023). Learning from crowds with robust support vector machines. *Science China Information Sciences*, 66(3), 132103..
68. *Gestión de Datos Autónoma*. Available online: <https://www.oracle.com/es/autonomous-database/>.
69. *SQL Developer*. Available online: <https://www.oracle.com/database/technologies/appdev/sqldeveloper-landing.html>.
70. *Oracle Data Miner* Available online: <https://www.oracle.com/big-data/technologies/dataminer/>

71. Chalapathy, R., Menon, A. K., & Chawla, S. (2018). *Anomaly detection using one-class neural networks*. *arXiv preprint arXiv:1802.06360*.
72. Oza, P., & Patel, V. M. (2018). *One-class convolutional neural network*. *IEEE Signal Processing Letters*, 26(2), 277-281.
73. Golpour, P., Ghayour-Mobarhan, M., Saki, A., Esmaily, H., Taghipour, A., Tajfard, M., ... & Ferns, G. A. (2020). *Comparison of support vector machine, naïve Bayes and logistic regression for assessing the necessity for coronary angiography*. *International journal of environmental research and public health*, 17(18), 6449.
74. Singh, S., Parmar, K. S., Makkhan, S. J. S., Kaur, J., Peshoria, S., & Kumar, J. (2020). *Study of ARIMA and least square support vector machine (LS-SVM) models for the prediction of SARS-CoV-2 confirmed cases in the most affected countries*. *Chaos, Solitons & Fractals*, 139, 110086.
75. Gupta, A. K., Singh, V., Mathur, P., & Travieso-Gonzalez, C. M. (2021). *Prediction of COVID-19 pandemic measuring criteria using support vector machine, prophet and linear regression models in Indian scenario*. *Journal of Interdisciplinary Mathematics*, 24(1), 89-108.
76. Byeon, H. (2021). *Predicting the severity of Parkinson's disease dementia by assessing the neuropsychiatric symptoms with an SVM regression model*. *International Journal of Environmental Research and Public Health*, 18(5), 2551.
77. Harimoorthy, K., & Thangavelu, M. (2023). *Retraction Note to: Multi-disease prediction model using improved*

SVM-radial bias technique in healthcare monitoring system.

78. Mughal, F. T. (2022). *AutoMAP–Smart Wearables for Depression: Toward Autonomous Mental Health Monitoring for the Elderly* (Doctoral dissertation).
79. Hastie, T. J. (Ed.). (2017). *Statistical models in S*. Routledge.
80. Scholkopf, B., & Smola, A. J. (2018). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press.
81. Motalo, K., Nojeem, L., Ewani, J., Opuiyo, A., & Browndi, I. (2023). Algebraic Multigrid and Cloud Computing: Enhancing Scalability and Performance. *International Journal of Technology and Scientific Research*, 12, 342-348.
82. Mohr, F., & van Rijn, J. N. (2023). Fast and informative model selection using learning curve cross-validation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

ANEXO

JAEN_V00_CUADRANTES_CP			
COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
CUADRANTE	NUMBER(5,0)	Yes	1
X_SI	NUMBER(19,11)	Yes	2
Y_SI	NUMBER(17,10)	Yes	3
X_SD	NUMBER(17,11)	Yes	4
Y_SD	NUMBER(19,10)	Yes	5
X_II	NUMBER(19,11)	Yes	6
Y_II	NUMBER(19,10)	Yes	7
X_ID	NUMBER(19,11)	Yes	8
Y_ID	NUMBER(19,10)	Yes	9
DIRECCIÓN1	VARCHAR2(512 BYTE)	Yes	10
NUMERO_DESDE	NUMBER(26,0)	Yes	11
NUMERO_HASTA	NUMBER(26,0)	Yes	12
POBLACION_TOTAL	NUMBER(10,0)	Yes	13
GRUPOS_EDAD_MENOR_16	NUMBER(10,0)	Yes	14
GRUPOS_EDAD_DE_16_A_64	NUMBER(10,0)	Yes	15
GRUPOS_EDAD_65_Y_MAS	NUMBER(10,0)	Yes	16
AFILIACION_SS	NUMBER(10,0)	Yes	17
T_CUENTA_PROPIA	NUMBER(10,0)	Yes	18
T_CUENTA_AJENA	NUMBER(10,0)	Yes	19
PENSIONISTAS	NUMBER(10,0)	Yes	20
CP	NUMBER(10,0)	Yes	21
CENTRO_DE_SALUD	VARCHAR2(26 BYTE)	Yes	22
DIRECCIÓN2	VARCHAR2(512 BYTE)	Yes	23

JAEN_EX_ESTACION_METEOROLOGICA

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
PROVINCIA_ID	VARCHAR2(2 BYTE)	No	1
ESTACION_ID	NUMBER(5,0)	No	2
ESTACION	VARCHAR2(100 BYTE)	No	3

JAEN_EX_INF_METEOROLOGICA

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
PROVINCIA_ID	VARCHAR2(2 BYTE)	No	1
ESTACION_ID	NUMBER	No	2
FECHA	DATE	No	3
DIA_JULIANO	NUMBER	Yes	4
T_MAX	NUMBER	Yes	5
HT_MAX	VARCHAR2(5 CHAR)	Yes	6
T_MIN	NUMBER	Yes	7
HT_MIN	VARCHAR2(5 CHAR)	Yes	8
T_MED	NUMBER	Yes	9
HUM_MAX	NUMBER	Yes	10
HUM_MIN	NUMBER	Yes	11
HUM_MED	NUMBER	Yes	12
VEL_VIENTO	NUMBER	Yes	13
DIR_VIENTO	NUMBER	Yes	14
RADIACION	NUMBER	Yes	15
PRECIPITACION	NUMBER	Yes	16
ETO	NUMBER	Yes	17

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
-------------	-----------	----------	-----------

JAEN_V00_CALLES_CP

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
TIPO_VIA	VARCHAR2(128 BYTE)	Yes	1
NOMBRE_VIA	VARCHAR2(512 BYTE)	Yes	2
NUMERACION	VARCHAR2(128 BYTE)	Yes	3
EXTREMO_INFERIOR	NUMBER(26,0)	Yes	4
EXTREMO_SUPERIOR	NUMBER(26,0)	Yes	5
CP	NUMBER(26,0)	Yes	6

JAEN_V02_CALLES_DIRDT

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
CALLE	VARCHAR2(256 BYTE)	Yes	1
NUMERO	VARCHAR2(26 BYTE)	Yes	2
CENTRO_SALUD	VARCHAR2(26 BYTE)	Yes	3
BASE	VARCHAR2(26 BYTE)	Yes	4

FESTIVOS_JAEN

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
FECHA_F	VARCHAR2(26 BYTE)	Yes	1
ANO_F	VARCHAR2(4 BYTE)	Yes	2
MES_F	VARCHAR2(2 BYTE)	Yes	3
DIA_F	VARCHAR2(2 BYTE)	Yes	4

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
DIA_SEMANA_F	VARCHAR2(20 BYTE)	Yes	5
SEMANA_MES_F	VARCHAR2(1 BYTE)	Yes	6

NUM_LAB_FES_INI

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	4
N_DIAS	NUMBER	Yes	5

JAEN_V00_DEMANDAS_AMP

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO_D	VARCHAR2(26 BYTE)	Yes	1
MES_D	VARCHAR2(6 BYTE)	Yes	2
DIA_D	VARCHAR2(6 BYTE)	Yes	3
IDPROVINCIA	VARCHAR2(26 BYTE)	Yes	4
FECHADEMANDA	VARCHAR2(26 BYTE)	Yes	5
IDDEMANDA	VARCHAR2(26 BYTE)	Yes	6
IDRECURSO	VARCHAR2(26 BYTE)	Yes	7
IDASISTENCIA	VARCHAR2(26 BYTE)	Yes	8
FECHALLAMADA	VARCHAR2(26 BYTE)	Yes	9
IDLLAMADA	VARCHAR2(26 BYTE)	Yes	10
IDLINEA	VARCHAR2(26 BYTE)	Yes	11

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
IDTIPOLINEA	VARCHAR2(26 BYTE)	Yes	12
IDALERTANTE	VARCHAR2(26 BYTE)	Yes	13
HORAENTRADA	VARCHAR2(26 BYTE)	Yes	14
HORACONTESTA	VARCHAR2(26 BYTE)	Yes	15
IDTIPODEMANDA	VARCHAR2(26 BYTE)	Yes	16
IDTIPODEMANDA1	VARCHAR2(26 BYTE)	Yes	17
IDTIPODEMANDA2	VARCHAR2(26 BYTE)	Yes	18
IDTIPODEMANDA3	VARCHAR2(26 BYTE)	Yes	19
FUENTECP	VARCHAR2(26 BYTE)	Yes	20
DIRIDPROVINCIARECURSO	VARCHAR2(256 BYTE)	Yes	21
DIRPEDANIARECURSO	VARCHAR2(256 BYTE)	Yes	22
DIRIDLOCALIDADRECURSO	VARCHAR2(256 BYTE)	Yes	23
DIRCALLE	VARCHAR2(256 BYTE)	Yes	24
DIRNUMERO	VARCHAR2(26 BYTE)	Yes	25
DIRPUERTA	VARCHAR2(26 BYTE)	Yes	26
DIRDT	VARCHAR2(26 BYTE)	Yes	27
DC	VARCHAR2(26 BYTE)	Yes	28
HORAACTIVACION	VARCHAR2(26 BYTE)	Yes	29
HORASALIDA	VARCHAR2(26 BYTE)	Yes	30
HORALLEGADA	VARCHAR2(26 BYTE)	Yes	31
HORACARGA	VARCHAR2(26 BYTE)	Yes	32
HORADESTINO	VARCHAR2(26 BYTE)	Yes	33
HORAOPERATIVA	VARCHAR2(26 BYTE)	Yes	34
HORADISPONIBLE	VARCHAR2(26 BYTE)	Yes	35
TIEMPOCOORDINACION	VARCHAR2(26 BYTE)	Yes	36
TIEMPOACTIVACION	VARCHAR2(26 BYTE)	Yes	37

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
TIEMPOATENCION	VARCHAR2(26 BYTE)	Yes	38
TIEMPORESPUESTA	VARCHAR2(26 BYTE)	Yes	39
IDMOTIVOEXCLUSION	VARCHAR2(26 BYTE)	Yes	40
IDUNIDADRECURSO	VARCHAR2(26 BYTE)	Yes	41
IDTIPORECURSO	VARCHAR2(26 BYTE)	Yes	42
IDEMPRESA	VARCHAR2(26 BYTE)	Yes	43
VEMERGENCIA	VARCHAR2(26 BYTE)	Yes	44
EDAD	VARCHAR2(26 BYTE)	Yes	45
SEXO	VARCHAR2(26 BYTE)	Yes	46
FNACIMIENTO	VARCHAR2(26 BYTE)	Yes	47
DIRIDPROVINCIA PACIENTE	VARCHAR2(256 BYTE)	Yes	48
DIRPEDANIAPACIENTE	VARCHAR2(256 BYTE)	Yes	49
DIRIDLOCALIDAD PACIENTE	VARCHAR2(256 BYTE)	Yes	50
DIRCALLE PACIENTE	VARCHAR2(256 BYTE)	Yes	51
DIRNUMERO PACIENTE	VARCHAR2(26 BYTE)	Yes	52
DIRPUERTA PACIENTE	VARCHAR2(26 BYTE)	Yes	53
JUICIOCLINICO1	VARCHAR2(26 BYTE)	Yes	54
JUICIOCLINICO2	VARCHAR2(26 BYTE)	Yes	55
JUICIOCLINICO3	VARCHAR2(26 BYTE)	Yes	56
IDCODIGORESOLUCION	VARCHAR2(26 BYTE)	Yes	57
IDFINANCIACION	VARCHAR2(26 BYTE)	Yes	58
DIRDESTINODT	VARCHAR2(256 BYTE)	Yes	59
SITUACION	VARCHAR2(26 BYTE)	Yes	60
SITUACIONDIRDT	VARCHAR2(10 BYTE)	Yes	61
IDCENTROINGRESO	VARCHAR2(256 BYTE)	Yes	62
IDBASE	VARCHAR2(26 BYTE)	Yes	63

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
HORALLAMADA	VARCHAR2(256 BYTE)	Yes	64
BASE	VARCHAR2(26 BYTE)	Yes	65
CENTRO	VARCHAR2(90 BYTE)	Yes	66
EMPRESA	VARCHAR2(26 BYTE)	Yes	67
HORADC	VARCHAR2(256 BYTE)	Yes	68
PRIORIDAD	VARCHAR2(5 BYTE)	Yes	69
ORDEN	VARCHAR2(256 BYTE)	Yes	70
LTR	VARCHAR2(50 BYTE)	Yes	71
DIRORIGENDT	VARCHAR2(256 BYTE)	Yes	72
PRIORIDADTIPORECURSO	VARCHAR2(5 BYTE)	Yes	73
IDTIPORECURSOAGR	VARCHAR2(26 BYTE)	Yes	74
TIEMPORECUPERACION	VARCHAR2(25 BYTE)	Yes	75
TIEMPODEMORA	VARCHAR2(25 BYTE)	Yes	76
IDTIPOLINEAUNICO	VARCHAR2(23 BYTE)	Yes	77
HORACUELGA	VARCHAR2(256 BYTE)	Yes	78
CADREC	VARCHAR2(256 BYTE)	Yes	79
IDESCENARIO	VARCHAR2(26 BYTE)	Yes	80
TIPOEDAD	VARCHAR2(6 BYTE)	Yes	81
IDMOTIVOEXCLUSION2	VARCHAR2(26 BYTE)	Yes	82
HORACREACIONRECURSO	VARCHAR2(26 BYTE)	Yes	83
HORASUENA	VARCHAR2(256 BYTE)	Yes	84
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	85
DIA_SEMANA	VARCHAR2(10 BYTE)	Yes	86
CP	NUMBER(5,0)	Yes	87

JAEN_V00_VALORES_X_CP

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
CPRO	NUMBER(4,0)	Yes	1
CMUN	NUMBER(2,0)	Yes	2
CP	NUMBER(5,0)	Yes	3
POBLACION_TOTAL	NUMBER(5,0)	Yes	4
HOMBRES	NUMBER(5,0)	Yes	5
MUJERES	NUMBER(5,0)	Yes	6
PERSONAS_MENOS_16	NUMBER(4,0)	Yes	7
PERSONAS_16_A_64	NUMBER(5,0)	Yes	8
PERSONAS_MAS_64	NUMBER(4,0)	Yes	9
HOMBRES_MENOS_16	NUMBER(4,0)	Yes	10
HOMBRES_16_A_64	NUMBER(5,0)	Yes	11
HOMBRES_MAS_64	NUMBER(4,0)	Yes	12
MUJERES_MENOS_16	NUMBER(4,0)	Yes	13
MUJERES_16_A_64	NUMBER(5,0)	Yes	14
MUJERES_MAS_64	NUMBER(4,0)	Yes	15
PERSONAS_SOLTERAS	NUMBER(5,0)	Yes	16
PERSONAS_CASADAS	NUMBER(5,0)	Yes	17
PERSONAS_SEPARADAS	NUMBER(3,0)	Yes	18
PERSONAS_DIVORCIADAS	NUMBER(3,0)	Yes	19
PERSONAS_VIUDAS	NUMBER(4,0)	Yes	20
PERSONAS_SOLTERAS_MENOS_16	NUMBER(4,0)	Yes	21
PERSONAS_SOLTERAS_16_A_64	NUMBER(4,0)	Yes	22
PERSONAS_SOLTERAS_MAS_64	NUMBER(3,0)	Yes	23
PERSONAS_CASADAS_MENOS_16	NUMBER(2,0)	Yes	24

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
PERSONAS_CASADAS_16_A_64	NUMBER(5,0)	Yes	25
PERSONAS_CASADAS_MAS_64	NUMBER(4,0)	Yes	26
PERSONAS_SEPARADAS_MENOS_16	NUMBER(1,0)	Yes	27
PERSONAS_SEPARADAS_16_A_64	NUMBER(3,0)	Yes	28
PERSONAS_SEPARADAS_MAS_64	NUMBER(2,0)	Yes	29
PERSONAS_DIVORCIADAS_MENOS_16	NUMBER(1,0)	Yes	30
PERSONAS_DIVORCIADAS_16_Y_64	NUMBER(3,0)	Yes	31
PERSONAS_DIVORCIADAS_MAS_64	NUMBER(2,0)	Yes	32
PERSONAS_VIUDAS_MENOS_16	NUMBER(2,0)	Yes	33
PERSONAS_VIUDAS_ENTRE_16_Y_64	NUMBER(3,0)	Yes	34
PERSONAS_VIUDAS_MAS_64	NUMBER(4,0)	Yes	35
PERSONAS_ANALFABETAS	NUMBER(3,0)	Yes	36
PERSONAS_SIN_ESTUDIOS	NUMBER(4,0)	Yes	37
PER_ESTUDIOS_PRIMER_GRADO	NUMBER(4,0)	Yes	38
PER_CON_ESTUDIOS_DE_2_GRADO	NUMBER(5,0)	Yes	39
PER_CON_ESTUDIOS_DE_3_GRADO	NUMBER(4,0)	Yes	40
SIN_DATOS_DE_ESTUD_MENOS_16	NUMBER(4,0)	Yes	41
HOMBRES_ANALFABETOS	NUMBER(2,0)	Yes	42
MUJERES_ANALFABETAS	NUMBER(3,0)	Yes	43
HOMBRES_SIN_ESTUDIOS	NUMBER(3,0)	Yes	44
MUJERES_SIN_ESTUDIOS	NUMBER(4,0)	Yes	45
HOMBRES_ESTUDIOS_PRIMER_GRADO	NUMBER(4,0)	Yes	46
MUJERES_ESTUDIOS_PRIMER_GRADO	NUMBER(4,0)	Yes	47
HOMBRES_ESTUDIOS_2_GRADO	NUMBER(4,0)	Yes	48
MUJERES_ESTUDIOS_2_GRADO	NUMBER(4,0)	Yes	49
HOMBRES_ESTUDIOS_TERCER_GRADO	NUMBER(4,0)	Yes	50

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
MUJERES_ESTUDIOS_TERCER_GRADO	NUMBER(4,0)	Yes	51
H_SIN_DATOS_DE_ESTUD_MEN_16	NUMBER(4,0)	Yes	52
M_SIN_DATOS_DE_ESTUD_MEN_16	NUMBER(4,0)	Yes	53
PERSONAS_16_Y_64_ANALFAB	NUMBER(3,0)	Yes	54
PERSONAS_MAS_64_ANALFABETAS	NUMBER(3,0)	Yes	55
PERSONAS_16_Y_64_SIN_ESTUDIOS	NUMBER(3,0)	Yes	56
PER_MAS_64_SIN_ESTUDIOS	NUMBER(4,0)	Yes	57
PER_16_Y_64_ESTUD_1ER_G	NUMBER(4,0)	Yes	58
PER_MAS_64_ESTUD_1ER_G	NUMBER(3,0)	Yes	59
PER_16_Y_64_ESTUD_2_G	NUMBER(5,0)	Yes	60
PER_MAS_64_ESTUD_2_G	NUMBER(4,0)	Yes	61
PER_16_Y_64_ESTUD_3ER_G	NUMBER(4,0)	Yes	62
PER_MAS_64_ESTUD_3ER_G	NUMBER(3,0)	Yes	63
PER_MENOS_16_SIN_INF_ESTUD	NUMBER(4,0)	Yes	64
TOTAL_VIVIENDAS	NUMBER(5,0)	Yes	65
VIVIENDAS_PRINCIPALES	NUMBER(5,0)	Yes	66
VIVIENDAS_SECUNDARIAS	NUMBER(4,0)	Yes	67
VIVIENDAS_VACIAS	NUMBER(4,0)	Yes	68
VIV_EN_PROP_POR_COMPRA_PAG	NUMBER(4,0)	Yes	69
VIV_EN_PROP_PAGOS_PEND	NUMBER(4,0)	Yes	70
VIV_EN_PROP_HERENCIA_O_DONA	NUMBER(3,0)	Yes	71
VIV_EN_ALQUILER	NUMBER(3,0)	Yes	72
VIV_CEDIDA_O_BAJO_PRECIO	NUMBER(3,0)	Yes	73
VIV_EN_OTRO_TIPO_DE_TENENCIA	NUMBER(3,0)	Yes	74
VIV_DE_MENOS_DE_30M2	NUMBER(1,0)	Yes	75
VIV_30_45_M2	NUMBER(4,0)	Yes	76

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
VIV_46_60_M2	NUMBER(6,0)	Yes	77
VIV_61_75_M2	NUMBER(6,0)	Yes	78
VIV_76_90_M2	NUMBER(6,0)	Yes	79
VIV_91_105_M2	NUMBER(6,0)	Yes	80
VIV_106_120_M2	NUMBER(5,0)	Yes	81
VIV_121_150_M2	NUMBER(5,0)	Yes	82
VIV_151_180_M2	NUMBER(5,0)	Yes	83
VIV_DE_MAS_DE_180_M2	NUMBER(5,0)	Yes	84
VIV_CON_1_HABITACION	NUMBER(2,0)	Yes	85
VIV_CON_2_HAB	NUMBER(2,0)	Yes	86
VIV_CON_3_HAB	NUMBER(3,0)	Yes	87
VIV_CON_4_HAB	NUMBER(4,0)	Yes	88
VIV_CON_5_HAB	NUMBER(4,0)	Yes	89
VIV_CON_6_HAB	NUMBER(4,0)	Yes	90
VIV_CON_7_HAB	NUMBER(3,0)	Yes	91
VIV_CON_8_HAB	NUMBER(2,0)	Yes	92
VIV_CON_9_O_MAS_HAB	NUMBER(2,0)	Yes	93
TOTAL_HOG	NUMBER(5,0)	Yes	94
HOG_DE_1_PERSONA	NUMBER(4,0)	Yes	95
HOG_DE_2_PER	NUMBER(4,0)	Yes	96
HOG_DE_3_PER	NUMBER(4,0)	Yes	97
HOG_DE_4_PER	NUMBER(4,0)	Yes	98
HOG_DE_5_PER	NUMBER(3,0)	Yes	99
HOG_DE_6_O_MAS_PER	NUMBER(3,0)	Yes	100

JAEN_V01_DEM_AMP_FES_AG_CTRL_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	4
TEMPERATURA_MAXIMA	NUMBER	Yes	5
TEMPERATURA_MINIMA	NUMBER	Yes	6
TEMPERATURA_MEDIA	NUMBER	Yes	7
HUMEDAD_MAXIMA	NUMBER	Yes	8
HUMEDAD_MINIMA	NUMBER	Yes	9
HUMEDAD_MEDIA	NUMBER	Yes	10
VELOCIDAD_VIENTO	NUMBER	Yes	11
RADIACION	NUMBER	Yes	12
PRECIPITACION	NUMBER	Yes	13
ETO	NUMBER	Yes	14
N_RECURSOS	NUMBER	Yes	15
N_DIAS	NUMBER	Yes	16
MEDIA_RECURSOS	NUMBER	Yes	17

JAEN_V01_DEM_AMP_FES_AG_DTMN_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	4

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
TEMPERATURA_MAXIMA	NUMBER	Yes	5
TEMPERATURA_MINIMA	NUMBER	Yes	6
TEMPERATURA_MEDIA	NUMBER	Yes	7
HUMEDAD_MAXIMA	NUMBER	Yes	8
HUMEDAD_MINIMA	NUMBER	Yes	9
HUMEDAD_MEDIA	NUMBER	Yes	10
VELOCIDAD_VIENTO	NUMBER	Yes	11
RADIACION	NUMBER	Yes	12
PRECIPITACION	NUMBER	Yes	13
ETO	NUMBER	Yes	14
N_RECURSOS	NUMBER	Yes	15
N_DIAS	NUMBER	Yes	16
MEDIA_RECURSOS	NUMBER	Yes	17

JAEN_V01_DEM_AMP_LAB_AG_CTRL_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	4
TEMPERATURA_MAXIMA	NUMBER	Yes	5
TEMPERATURA_MINIMA	NUMBER	Yes	6
TEMPERATURA_MEDIA	NUMBER	Yes	7
HUMEDAD_MAXIMA	NUMBER	Yes	8
HUMEDAD_MINIMA	NUMBER	Yes	9

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
HUMEDAD_MEDIA	NUMBER	Yes	10
VELOCIDAD_VIENTO	NUMBER	Yes	11
RADIACION	NUMBER	Yes	12
PRECIPITACION	NUMBER	Yes	13
ETO	NUMBER	Yes	14
N_RECURSOS	NUMBER	Yes	15
N_DIAS	NUMBER	Yes	16
MEDIA_RECURSOS	NUMBER	Yes	17

JAEN_V01_DEM_AMP_LAB_AG_DTMN_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	4
TEMPERATURA_MAXIMA	NUMBER	Yes	5
TEMPERATURA_MINIMA	NUMBER	Yes	6
TEMPERATURA_MEDIA	NUMBER	Yes	7
HUMEDAD_MAXIMA	NUMBER	Yes	8
HUMEDAD_MINIMA	NUMBER	Yes	9
HUMEDAD_MEDIA	NUMBER	Yes	10
VELOCIDAD_VIENTO	NUMBER	Yes	11
RADIACION	NUMBER	Yes	12
PRECIPITACION	NUMBER	Yes	13
ETO	NUMBER	Yes	14

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
N_RECURSOS	NUMBER	Yes	15
N_DIAS	NUMBER	Yes	16
MEDIA_RECURSOS	NUMBER	Yes	17

JAEN_V01_DEMANDAS_A04_MES_CTL

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
TEMPERATURA_MAXIMA	NUMBER	Yes	3
TEMPERATURA_MINIMA	NUMBER	Yes	4
TEMPERATURA_MEDIA	NUMBER	Yes	5
HUMEDAD_MAXIMA	NUMBER	Yes	6
HUMEDAD_MINIMA	NUMBER	Yes	7
HUMEDAD_MEDIA	NUMBER	Yes	8
VELOCIDAD_VIENTO	NUMBER	Yes	9
RADIACION	NUMBER	Yes	10
PRECIPITACION	NUMBER	Yes	11
ETO	NUMBER	Yes	12
N_RECURSOS	NUMBER	Yes	13

JAEN_V01_DEMANDAS_A04_MES_DTM

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
TEMPERATURA_MAXIMA	NUMBER	Yes	3
TEMPERATURA_MINIMA	NUMBER	Yes	4
TEMPERATURA_MEDIA	NUMBER	Yes	5
HUMEDAD_MAXIMA	NUMBER	Yes	6
HUMEDAD_MINIMA	NUMBER	Yes	7
HUMEDAD_MEDIA	NUMBER	Yes	8
VELOCIDAD_VIENTO	NUMBER	Yes	9
RADIACION	NUMBER	Yes	10
PRECIPITACION	NUMBER	Yes	11
ETO	NUMBER	Yes	12
N_RECURSOS	NUMBER	Yes	13

JAEN_V02_DEMANDA_AG_CONTROL

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
FECHA	VARCHAR2(8 BYTE)	Yes	1
ANO	VARCHAR2(4 BYTE)	Yes	2
MES	VARCHAR2(40 BYTE)	Yes	3
DIA_SEMANA	VARCHAR2(36 BYTE)	Yes	4
DIA_FESTIVO	VARCHAR2(9 BYTE)	Yes	5
DIA_MES	VARCHAR2(2 BYTE)	Yes	6
DIA_ANO	VARCHAR2(3 BYTE)	Yes	7
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	8
SEMANA_ANO	VARCHAR2(2 BYTE)	Yes	9
DEMARCAACION	VARCHAR2(26 BYTE)	Yes	10
TIPO_RECURSO	VARCHAR2(26 BYTE)	Yes	11

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
TEMPERATURA_MAXIMA	NUMBER	Yes	12
TEMPERATURA_MINIMA	NUMBER	Yes	13
TEMPERATURA_MEDIA	NUMBER	Yes	14
HUMEDAD_MAXIMA	NUMBER	Yes	15
HUMEDAD_MINIMA	NUMBER	Yes	16
HUMEDAD_MEDIA	NUMBER	Yes	17
VELOCIDAD_VIENTO	NUMBER	Yes	18
RADIACION	NUMBER	Yes	19
PRECIPITACION	NUMBER	Yes	20
ETO	NUMBER	Yes	21
N_RECURSOS	NUMBER	Yes	22

JAEN_V02_DEMANDA_AG_DATAMINING

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
FECHA	VARCHAR2(8 BYTE)	Yes	1
ANO	VARCHAR2(4 BYTE)	Yes	2
MES	VARCHAR2(40 BYTE)	Yes	3
DIA_SEMANA	VARCHAR2(36 BYTE)	Yes	4
DIA_FESTIVO	VARCHAR2(9 BYTE)	Yes	5
DIA_MES	VARCHAR2(2 BYTE)	Yes	6
DIA_ANO	VARCHAR2(3 BYTE)	Yes	7
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	8
SEMANA_ANO	VARCHAR2(2 BYTE)	Yes	9
DEMARCAION	VARCHAR2(26 BYTE)	Yes	10
TIPO_RECURSO	VARCHAR2(26 BYTE)	Yes	11

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
TEMPERATURA_MAXIMA	NUMBER	Yes	12
TEMPERATURA_MINIMA	NUMBER	Yes	13
TEMPERATURA_MEDIA	NUMBER	Yes	14
HUMEDAD_MAXIMA	NUMBER	Yes	15
HUMEDAD_MINIMA	NUMBER	Yes	16
HUMEDAD_MEDIA	NUMBER	Yes	17
VELOCIDAD_VIENTO	NUMBER	Yes	18
RADIACION	NUMBER	Yes	19
PRECIPITACION	NUMBER	Yes	20
ETO	NUMBER	Yes	21
N_RECURSOS	NUMBER	Yes	22

JAEN_V02_DEMANDAS_AMP_F

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO_D	VARCHAR2(26 BYTE)	Yes	1
MES_D	VARCHAR2(6 BYTE)	Yes	2
DIA_D	VARCHAR2(6 BYTE)	Yes	3
IDPROVINCIA	VARCHAR2(26 BYTE)	Yes	4
FECHADEMANDA	VARCHAR2(26 BYTE)	Yes	5
IDDEMANDA	VARCHAR2(26 BYTE)	Yes	6
IDRECURSO	VARCHAR2(26 BYTE)	Yes	7
IDASISTENCIA	VARCHAR2(26 BYTE)	Yes	8
FECHALLAMADA	VARCHAR2(26 BYTE)	Yes	9
IDLLAMADA	VARCHAR2(26 BYTE)	Yes	10
IDLINEA	VARCHAR2(26 BYTE)	Yes	11

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
IDTIPOLINEA	VARCHAR2(26 BYTE)	Yes	12
IDALERTANTE	VARCHAR2(26 BYTE)	Yes	13
HORAENTRADA	VARCHAR2(26 BYTE)	Yes	14
HORACONTESTA	VARCHAR2(26 BYTE)	Yes	15
IDTIPODEMANDA	VARCHAR2(26 BYTE)	Yes	16
IDTIPODEMANDA1	VARCHAR2(26 BYTE)	Yes	17
IDTIPODEMANDA2	VARCHAR2(26 BYTE)	Yes	18
IDTIPODEMANDA3	VARCHAR2(26 BYTE)	Yes	19
FUENTECP	VARCHAR2(26 BYTE)	Yes	20
DIRIDPROVINCIARECURSO	VARCHAR2(256 BYTE)	Yes	21
DIRPEDANIARECURSO	VARCHAR2(256 BYTE)	Yes	22
DIRIDLOCALIDADRECURSO	VARCHAR2(256 BYTE)	Yes	23
DIRCALLE	VARCHAR2(256 BYTE)	Yes	24
DIRNUMERO	VARCHAR2(26 BYTE)	Yes	25
DIRPUERTA	VARCHAR2(26 BYTE)	Yes	26
DIRDT	VARCHAR2(26 BYTE)	Yes	27
DC	VARCHAR2(26 BYTE)	Yes	28
HORAACTIVACION	VARCHAR2(26 BYTE)	Yes	29
HORASALIDA	VARCHAR2(26 BYTE)	Yes	30
HORALLEGADA	VARCHAR2(26 BYTE)	Yes	31
HORACARGA	VARCHAR2(26 BYTE)	Yes	32
HORADESTINO	VARCHAR2(26 BYTE)	Yes	33
HORAOPERATIVA	VARCHAR2(26 BYTE)	Yes	34
HORADISPONIBLE	VARCHAR2(26 BYTE)	Yes	35
TIEMPOCOORDINACION	VARCHAR2(26 BYTE)	Yes	36
TIEMPOACTIVACION	VARCHAR2(26 BYTE)	Yes	37

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
TIEMPOATENCION	VARCHAR2(26 BYTE)	Yes	38
TIEMPORESPUESTA	VARCHAR2(26 BYTE)	Yes	39
IDMOTIVOEXCLUSION	VARCHAR2(26 BYTE)	Yes	40
IDUNIDADRECURSO	VARCHAR2(26 BYTE)	Yes	41
IDTIPORECURSO	VARCHAR2(26 BYTE)	Yes	42
IDEMPRESA	VARCHAR2(26 BYTE)	Yes	43
VEMERGENCIA	VARCHAR2(26 BYTE)	Yes	44
EDAD	VARCHAR2(26 BYTE)	Yes	45
SEXO	VARCHAR2(26 BYTE)	Yes	46
FNACIMIENTO	VARCHAR2(26 BYTE)	Yes	47
DIRIDPROVINCIA PACIENTE	VARCHAR2(256 BYTE)	Yes	48
DIRPEDANIAPACIENTE	VARCHAR2(256 BYTE)	Yes	49
DIRIDLOCALIDADPACIENTE	VARCHAR2(256 BYTE)	Yes	50
DIRCALLEPACIENTE	VARCHAR2(256 BYTE)	Yes	51
DIRNUMEROPACIENTE	VARCHAR2(26 BYTE)	Yes	52
DIRPUERTAPACIENTE	VARCHAR2(26 BYTE)	Yes	53
JUICIOCLINICO1	VARCHAR2(26 BYTE)	Yes	54
JUICIOCLINICO2	VARCHAR2(26 BYTE)	Yes	55
JUICIOCLINICO3	VARCHAR2(26 BYTE)	Yes	56
IDCODIGORESOLUCION	VARCHAR2(26 BYTE)	Yes	57
IDFINANCIACION	VARCHAR2(26 BYTE)	Yes	58
DIRDESTINODT	VARCHAR2(256 BYTE)	Yes	59
SITUACION	VARCHAR2(26 BYTE)	Yes	60
SITUACIONDIRDT	VARCHAR2(10 BYTE)	Yes	61
IDCENTROINGRESO	VARCHAR2(256 BYTE)	Yes	62
IDBASE	VARCHAR2(26 BYTE)	Yes	63

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
HORALLAMADA	VARCHAR2(256 BYTE)	Yes	64
BASE	VARCHAR2(26 BYTE)	Yes	65
CENTRO	VARCHAR2(90 BYTE)	Yes	66
EMPRESA	VARCHAR2(26 BYTE)	Yes	67
HORADC	VARCHAR2(256 BYTE)	Yes	68
PRIORIDAD	VARCHAR2(5 BYTE)	Yes	69
ORDEN	VARCHAR2(256 BYTE)	Yes	70
LTR	VARCHAR2(50 BYTE)	Yes	71
DIRORIGENDT	VARCHAR2(256 BYTE)	Yes	72
PRIORIDADTIPORECURSO	VARCHAR2(5 BYTE)	Yes	73
IDTIPORECURSOAGR	VARCHAR2(26 BYTE)	Yes	74
TIEMPORECUPERACION	VARCHAR2(25 BYTE)	Yes	75
TIEMPODEMORA	VARCHAR2(25 BYTE)	Yes	76
IDTIPOLINEAUNICO	VARCHAR2(23 BYTE)	Yes	77
HORACUELGA	VARCHAR2(256 BYTE)	Yes	78
CADREC	VARCHAR2(256 BYTE)	Yes	79
IDESCENARIO	VARCHAR2(26 BYTE)	Yes	80
TIPOEDAD	VARCHAR2(6 BYTE)	Yes	81
IDMOTIVOEXCLUSION2	VARCHAR2(26 BYTE)	Yes	82
HORACREACIONRECURSO	VARCHAR2(26 BYTE)	Yes	83
HORASUENA	VARCHAR2(256 BYTE)	Yes	84
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	85
DIA_SEMANA	VARCHAR2(10 BYTE)	Yes	86
MT	VARCHAR2(26 BYTE)	Yes	87

JAEN_V02_DEMANDAS_DIA_SEMANA

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
DIA_SEMANA	VARCHAR2(10 BYTE)	Yes	2
TEMPERATURA_MAXIMA	NUMBER	Yes	3
TEMPERATURA_MINIMA	NUMBER	Yes	4
TEMPERATURA_MEDIA	NUMBER	Yes	5
HUMEDAD_MAXIMA	NUMBER	Yes	6
HUMEDAD_MINIMA	NUMBER	Yes	7
HUMEDAD_MEDIA	NUMBER	Yes	8
VELOCIDAD_VIENTO	NUMBER	Yes	9
RADIACION	NUMBER	Yes	10
PRECIPITACION	NUMBER	Yes	11
ETO	NUMBER	Yes	12
N_RECURSOS	NUMBER	Yes	13

JAEN_V10_DEMANDA_FES_AG_CTRL_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TEMPERATURA_MAXIMA	NUMBER	Yes	4
TEMPERATURA_MINIMA	NUMBER	Yes	5
TEMPERATURA_MEDIA	NUMBER	Yes	6
HUMEDAD_MAXIMA	NUMBER	Yes	7
HUMEDAD_MINIMA	NUMBER	Yes	8

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
HUMEDAD_MEDIA	NUMBER	Yes	9
VELOCIDAD_VIENTO	NUMBER	Yes	10
RADIACION	NUMBER	Yes	11
PRECIPITACION	NUMBER	Yes	12
ETO	NUMBER	Yes	13
N_RECURSOS	NUMBER	Yes	14

JAEN_V10_DEMANDA_FES_AG_DTMN_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TEMPERATURA_MAXIMA	NUMBER	Yes	4
TEMPERATURA_MINIMA	NUMBER	Yes	5
TEMPERATURA_MEDIA	NUMBER	Yes	6
HUMEDAD_MAXIMA	NUMBER	Yes	7
HUMEDAD_MINIMA	NUMBER	Yes	8
HUMEDAD_MEDIA	NUMBER	Yes	9
VELOCIDAD_VIENTO	NUMBER	Yes	10
RADIACION	NUMBER	Yes	11
PRECIPITACION	NUMBER	Yes	12
ETO	NUMBER	Yes	13
N_RECURSOS	NUMBER	Yes	14

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TEMPERATURA_MAXIMA	NUMBER	Yes	4
TEMPERATURA_MINIMA	NUMBER	Yes	5
TEMPERATURA_MEDIA	NUMBER	Yes	6
HUMEDAD_MAXIMA	NUMBER	Yes	7
HUMEDAD_MINIMA	NUMBER	Yes	8
HUMEDAD_MEDIA	NUMBER	Yes	9
VELOCIDAD_VIENTO	NUMBER	Yes	10
RADIACION	NUMBER	Yes	11
PRECIPITACION	NUMBER	Yes	12
ETO	NUMBER	Yes	13
N_RECURSOS	NUMBER	Yes	14

JAEN_V10_DEMANDA_LAB_AG_DTMN_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TEMPERATURA_MAXIMA	NUMBER	Yes	4
TEMPERATURA_MINIMA	NUMBER	Yes	5
TEMPERATURA_MEDIA	NUMBER	Yes	6
HUMEDAD_MAXIMA	NUMBER	Yes	7

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
HUMEDAD_MINIMA	NUMBER	Yes	8
HUMEDAD_MEDIA	NUMBER	Yes	9
VELOCIDAD_VIENTO	NUMBER	Yes	10
RADIACION	NUMBER	Yes	11
PRECIPITACION	NUMBER	Yes	12
ETO	NUMBER	Yes	13
N_RECURSOS	NUMBER	Yes	14

JAEN_V100_DEM_XREC_AG_CTL

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
FECHA	VARCHAR2(8 BYTE)	Yes	1
ANO	VARCHAR2(4 BYTE)	Yes	2
MES	VARCHAR2(40 BYTE)	Yes	3
DIA_SEMANA	VARCHAR2(36 BYTE)	Yes	4
DIA_MES	VARCHAR2(2 BYTE)	Yes	5
DIA_ANO	VARCHAR2(3 BYTE)	Yes	6
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	7
SEMANA_ANO	VARCHAR2(2 BYTE)	Yes	8
TIPO_RECURSO	VARCHAR2(26 BYTE)	Yes	9
TEMPERATURA_MAXIMA	NUMBER	Yes	10
TEMPERATURA_MINIMA	NUMBER	Yes	11
TEMPERATURA_MEDIA	NUMBER	Yes	12
HUMEDAD_MAXIMA	NUMBER	Yes	13
HUMEDAD_MINIMA	NUMBER	Yes	14
HUMEDAD_MEDIA	NUMBER	Yes	15

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
VELOCIDAD_VIENTO	NUMBER	Yes	16
RADIACION	NUMBER	Yes	17
PRECIPITACION	NUMBER	Yes	18
ETO	NUMBER	Yes	19
N_RECURSOS	NUMBER	Yes	20
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	21

JAEN_V100_DEM_XREC_AG_DTM

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
FECHA	VARCHAR2(8 BYTE)	Yes	1
ANO	VARCHAR2(4 BYTE)	Yes	2
MES	VARCHAR2(40 BYTE)	Yes	3
DIA_SEMANA	VARCHAR2(36 BYTE)	Yes	4
DIA_MES	VARCHAR2(2 BYTE)	Yes	5
DIA_ANO	VARCHAR2(3 BYTE)	Yes	6
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	7
SEMANA_ANO	VARCHAR2(2 BYTE)	Yes	8
TIPO_RECURSO	VARCHAR2(26 BYTE)	Yes	9
TEMPERATURA_MAXIMA	NUMBER	Yes	10
TEMPERATURA_MINIMA	NUMBER	Yes	11
TEMPERATURA_MEDIA	NUMBER	Yes	12
HUMEDAD_MAXIMA	NUMBER	Yes	13
HUMEDAD_MINIMA	NUMBER	Yes	14
HUMEDAD_MEDIA	NUMBER	Yes	15
VELOCIDAD_VIENTO	NUMBER	Yes	16

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
RADIACION	NUMBER	Yes	17
PRECIPITACION	NUMBER	Yes	18
ETO	NUMBER	Yes	19
N_RECursos	NUMBER	Yes	20
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	21

JAEN_V101_DEM_REC_AG_CTL

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
FECHA	VARCHAR2(8 BYTE)	Yes	1
ANO	VARCHAR2(4 BYTE)	Yes	2
MES	VARCHAR2(40 BYTE)	Yes	3
DIA_SEMANA	VARCHAR2(36 BYTE)	Yes	4
DIA_MES	VARCHAR2(2 BYTE)	Yes	5
DIA_ANO	VARCHAR2(3 BYTE)	Yes	6
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	7
SEMANA_ANO	VARCHAR2(2 BYTE)	Yes	8
TEMPERATURA_MAXIMA	NUMBER	Yes	9
TEMPERATURA_MINIMA	NUMBER	Yes	10
TEMPERATURA_MEDIA	NUMBER	Yes	11
HUMEDAD_MAXIMA	NUMBER	Yes	12
HUMEDAD_MINIMA	NUMBER	Yes	13
HUMEDAD_MEDIA	NUMBER	Yes	14
VELOCIDAD_VIENTO	NUMBER	Yes	15
RADIACION	NUMBER	Yes	16
PRECIPITACION	NUMBER	Yes	17

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ETO	NUMBER	Yes	18
N_RECURSOS	NUMBER	Yes	19
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	20

JAEN_V101_DEM_REC_AG_DTM

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
FECHA	VARCHAR2(8 BYTE)	Yes	1
ANO	VARCHAR2(4 BYTE)	Yes	2
MES	VARCHAR2(40 BYTE)	Yes	3
DIA_SEMANA	VARCHAR2(36 BYTE)	Yes	4
DIA_MES	VARCHAR2(2 BYTE)	Yes	5
DIA_ANO	VARCHAR2(3 BYTE)	Yes	6
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	7
SEMANA_ANO	VARCHAR2(2 BYTE)	Yes	8
TEMPERATURA_MAXIMA	NUMBER	Yes	9
TEMPERATURA_MINIMA	NUMBER	Yes	10
TEMPERATURA_MEDIA	NUMBER	Yes	11
HUMEDAD_MAXIMA	NUMBER	Yes	12
HUMEDAD_MINIMA	NUMBER	Yes	13
HUMEDAD_MEDIA	NUMBER	Yes	14
VELOCIDAD_VIENTO	NUMBER	Yes	15
RADIACION	NUMBER	Yes	16
PRECIPITACION	NUMBER	Yes	17
ETO	NUMBER	Yes	18
N_RECURSOS	NUMBER	Yes	19

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	20

JAEN_V102_DEM_A04_REC_AG_CTL

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
FECHA	VARCHAR2(8 BYTE)	Yes	1
ANO	VARCHAR2(4 BYTE)	Yes	2
MES	VARCHAR2(40 BYTE)	Yes	3
DIA_SEMANA	VARCHAR2(36 BYTE)	Yes	4
DIA_MES	VARCHAR2(2 BYTE)	Yes	5
DIA_ANO	VARCHAR2(3 BYTE)	Yes	6
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	7
SEMANA_ANO	VARCHAR2(2 BYTE)	Yes	8
TEMPERATURA_MAXIMA	NUMBER	Yes	9
TEMPERATURA_MINIMA	NUMBER	Yes	10
TEMPERATURA_MEDIA	NUMBER	Yes	11
HUMEDAD_MAXIMA	NUMBER	Yes	12
HUMEDAD_MINIMA	NUMBER	Yes	13
HUMEDAD_MEDIA	NUMBER	Yes	14
VELOCIDAD_VIENTO	NUMBER	Yes	15
RADIACION	NUMBER	Yes	16
PRECIPITACION	NUMBER	Yes	17
ETO	NUMBER	Yes	18
N_RECURSOS	NUMBER	Yes	19
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	20

JAEN_V102_DEM_A04_REC_AG_DTM

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
FECHA	VARCHAR2(8 BYTE)	Yes	1
ANO	VARCHAR2(4 BYTE)	Yes	2
MES	VARCHAR2(40 BYTE)	Yes	3
DIA_SEMANA	VARCHAR2(36 BYTE)	Yes	4
DIA_MES	VARCHAR2(2 BYTE)	Yes	5
DIA_ANO	VARCHAR2(3 BYTE)	Yes	6
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	7
SEMANA_ANO	VARCHAR2(2 BYTE)	Yes	8
TEMPERATURA_MAXIMA	NUMBER	Yes	9
TEMPERATURA_MINIMA	NUMBER	Yes	10
TEMPERATURA_MEDIA	NUMBER	Yes	11
HUMEDAD_MAXIMA	NUMBER	Yes	12
HUMEDAD_MINIMA	NUMBER	Yes	13
HUMEDAD_MEDIA	NUMBER	Yes	14
VELOCIDAD_VIENTO	NUMBER	Yes	15
RADIACION	NUMBER	Yes	16
PRECIPITACION	NUMBER	Yes	17
ETO	NUMBER	Yes	18
N_RECURSOS	NUMBER	Yes	19
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	20

JAEN_V103_DEM_A04_XREC_AG_CTL

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
FECHA	VARCHAR2(8 BYTE)	Yes	1

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	2
MES	VARCHAR2(40 BYTE)	Yes	3
DIA_SEMANA	VARCHAR2(36 BYTE)	Yes	4
DIA_MES	VARCHAR2(2 BYTE)	Yes	5
DIA_ANO	VARCHAR2(3 BYTE)	Yes	6
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	7
SEMANA_ANO	VARCHAR2(2 BYTE)	Yes	8
TIPO_RECURSO	VARCHAR2(26 BYTE)	Yes	9
TEMPERATURA_MAXIMA	NUMBER	Yes	10
TEMPERATURA_MINIMA	NUMBER	Yes	11
TEMPERATURA_MEDIA	NUMBER	Yes	12
HUMEDAD_MAXIMA	NUMBER	Yes	13
HUMEDAD_MINIMA	NUMBER	Yes	14
HUMEDAD_MEDIA	NUMBER	Yes	15
VELOCIDAD_VIENTO	NUMBER	Yes	16
RADIACION	NUMBER	Yes	17
PRECIPITACION	NUMBER	Yes	18
ETO	NUMBER	Yes	19
N_RECURSOS	NUMBER	Yes	20
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	21

JAEN_V103_DEM_A04_XREC_AG_DTM

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
FECHA	VARCHAR2(8 BYTE)	Yes	1
ANO	VARCHAR2(4 BYTE)	Yes	2

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
MES	VARCHAR2(40 BYTE)	Yes	3
DIA_SEMANA	VARCHAR2(36 BYTE)	Yes	4
DIA_MES	VARCHAR2(2 BYTE)	Yes	5
DIA_AÑO	VARCHAR2(3 BYTE)	Yes	6
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	7
SEMANA_AÑO	VARCHAR2(2 BYTE)	Yes	8
TIPO_RECURSO	VARCHAR2(26 BYTE)	Yes	9
TEMPERATURA_MAXIMA	NUMBER	Yes	10
TEMPERATURA_MINIMA	NUMBER	Yes	11
TEMPERATURA_MEDIA	NUMBER	Yes	12
HUMEDAD_MAXIMA	NUMBER	Yes	13
HUMEDAD_MINIMA	NUMBER	Yes	14
HUMEDAD_MEDIA	NUMBER	Yes	15
VELOCIDAD_VIENTO	NUMBER	Yes	16
RADIACION	NUMBER	Yes	17
PRECIPITACION	NUMBER	Yes	18
ETO	NUMBER	Yes	19
N_RECURSOS	NUMBER	Yes	20
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	21

JAEN_V104_DEM_A04_REC_AG_CTL

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
FECHA	VARCHAR2(8 BYTE)	Yes	1
AÑO	VARCHAR2(4 BYTE)	Yes	2
MES	VARCHAR2(40 BYTE)	Yes	3

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
DIA_SEMANA	VARCHAR2(36 BYTE)	Yes	4
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	5
DIA_MES	VARCHAR2(2 BYTE)	Yes	6
DIA_ANO	VARCHAR2(3 BYTE)	Yes	7
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	8
SEMANA_ANO	VARCHAR2(2 BYTE)	Yes	9
TEMPERATURA_MAXIMA	NUMBER	Yes	10
TEMPERATURA_MINIMA	NUMBER	Yes	11
TEMPERATURA_MEDIA	NUMBER	Yes	12
HUMEDAD_MAXIMA	NUMBER	Yes	13
HUMEDAD_MINIMA	NUMBER	Yes	14
HUMEDAD_MEDIA	NUMBER	Yes	15
VELOCIDAD_VIENTO	NUMBER	Yes	16
RADIACION	NUMBER	Yes	17
PRECIPITACION	NUMBER	Yes	18
ETO	NUMBER	Yes	19
N_RECURSOS	NUMBER	Yes	20

JAEN_V104_DEM_A04_REC_AG_DTM

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
FECHA	VARCHAR2(8 BYTE)	Yes	1
ANO	VARCHAR2(4 BYTE)	Yes	2
MES	VARCHAR2(40 BYTE)	Yes	3
DIA_SEMANA	VARCHAR2(36 BYTE)	Yes	4
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	5

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
DIA_MES	VARCHAR2(2 BYTE)	Yes	6
DIA_ANO	VARCHAR2(3 BYTE)	Yes	7
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	8
SEMANA_ANO	VARCHAR2(2 BYTE)	Yes	9
TEMPERATURA_MAXIMA	NUMBER	Yes	10
TEMPERATURA_MINIMA	NUMBER	Yes	11
TEMPERATURA_MEDIA	NUMBER	Yes	12
HUMEDAD_MAXIMA	NUMBER	Yes	13
HUMEDAD_MINIMA	NUMBER	Yes	14
HUMEDAD_MEDIA	NUMBER	Yes	15
VELOCIDAD_VIENTO	NUMBER	Yes	16
RADIACION	NUMBER	Yes	17
PRECIPITACION	NUMBER	Yes	18
ETO	NUMBER	Yes	19
N_RECURSOS	NUMBER	Yes	20

JAEN_V11_DEMANDA_FES_AG_CTRL_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TEMPERATURA_MAXIMA	NUMBER	Yes	4
TEMPERATURA_MINIMA	NUMBER	Yes	5
TEMPERATURA_MEDIA	NUMBER	Yes	6
HUMEDAD_MAXIMA	NUMBER	Yes	7

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
HUMEDAD_MINIMA	NUMBER	Yes	8
HUMEDAD_MEDIA	NUMBER	Yes	9
VELOCIDAD_VIENTO	NUMBER	Yes	10
RADIACION	NUMBER	Yes	11
PRECIPITACION	NUMBER	Yes	12
ETO	NUMBER	Yes	13
N_RECURSOS	NUMBER	Yes	14

JAEN_V11_DEMANDA_FES_AG_CTRL_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TEMPERATURA_MAXIMA	NUMBER	Yes	4
TEMPERATURA_MINIMA	NUMBER	Yes	5
TEMPERATURA_MEDIA	NUMBER	Yes	6
HUMEDAD_MAXIMA	NUMBER	Yes	7
HUMEDAD_MINIMA	NUMBER	Yes	8
HUMEDAD_MEDIA	NUMBER	Yes	9
VELOCIDAD_VIENTO	NUMBER	Yes	10
RADIACION	NUMBER	Yes	11
PRECIPITACION	NUMBER	Yes	12
ETO	NUMBER	Yes	13
N_RECURSOS	NUMBER	Yes	14

JAEN_V11_DEMANDA_FES_AG_DTMN_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TEMPERATURA_MAXIMA	NUMBER	Yes	4
TEMPERATURA_MINIMA	NUMBER	Yes	5
TEMPERATURA_MEDIA	NUMBER	Yes	6
HUMEDAD_MAXIMA	NUMBER	Yes	7
HUMEDAD_MINIMA	NUMBER	Yes	8
HUMEDAD_MEDIA	NUMBER	Yes	9
VELOCIDAD_VIENTO	NUMBER	Yes	10
RADIACION	NUMBER	Yes	11
PRECIPITACION	NUMBER	Yes	12
ETO	NUMBER	Yes	13
N_RECursos	NUMBER	Yes	14

JAEN_V11_DEMANDA_LAB_AG_CTRL_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TEMPERATURA_MAXIMA	NUMBER	Yes	4
TEMPERATURA_MINIMA	NUMBER	Yes	5
TEMPERATURA_MEDIA	NUMBER	Yes	6
HUMEDAD_MAXIMA	NUMBER	Yes	7

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
HUMEDAD_MINIMA	NUMBER	Yes	8
HUMEDAD_MEDIA	NUMBER	Yes	9
VELOCIDAD_VIENTO	NUMBER	Yes	10
RADIACION	NUMBER	Yes	11
PRECIPITACION	NUMBER	Yes	12
ETO	NUMBER	Yes	13
N_RECURSOS	NUMBER	Yes	14

JAEN_V11_DEMANDA_LAB_AG_DTMN_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TEMPERATURA_MAXIMA	NUMBER	Yes	4
TEMPERATURA_MINIMA	NUMBER	Yes	5
TEMPERATURA_MEDIA	NUMBER	Yes	6
HUMEDAD_MAXIMA	NUMBER	Yes	7
HUMEDAD_MINIMA	NUMBER	Yes	8
HUMEDAD_MEDIA	NUMBER	Yes	9
VELOCIDAD_VIENTO	NUMBER	Yes	10
RADIACION	NUMBER	Yes	11
PRECIPITACION	NUMBER	Yes	12
ETO	NUMBER	Yes	13
N_RECURSOS	NUMBER	Yes	14

JAEN_V12_DEM_AMP_FES_AG_CTRL_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	4
TEMPERATURA_MAXIMA	NUMBER	Yes	5
TEMPERATURA_MINIMA	NUMBER	Yes	6
TEMPERATURA_MEDIA	NUMBER	Yes	7
HUMEDAD_MAXIMA	NUMBER	Yes	8
HUMEDAD_MINIMA	NUMBER	Yes	9
HUMEDAD_MEDIA	NUMBER	Yes	10
VELOCIDAD_VIENTO	NUMBER	Yes	11
RADIACION	NUMBER	Yes	12
PRECIPITACION	NUMBER	Yes	13
ETO	NUMBER	Yes	14
N_RECURSOS	NUMBER	Yes	15
N_DIAS	NUMBER	Yes	16
MEDIA_RECURSOS	NUMBER	Yes	17

JAEN_V12_DEM_AMP_FES_AG_DTMN_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	4

Predicción de las patologías más comunes y emergencias sanitarias basados en variables espaciales mediante el uso de la minería de datos

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
TEMPERATURA_MAXIMA	NUMBER	Yes	5
TEMPERATURA_MINIMA	NUMBER	Yes	6
TEMPERATURA_MEDIA	NUMBER	Yes	7
HUMEDAD_MAXIMA	NUMBER	Yes	8
HUMEDAD_MINIMA	NUMBER	Yes	9
HUMEDAD_MEDIA	NUMBER	Yes	10
VELOCIDAD_VIENTO	NUMBER	Yes	11
RADIACION	NUMBER	Yes	12
PRECIPITACION	NUMBER	Yes	13
ETO	NUMBER	Yes	14
N_RECURSOS	NUMBER	Yes	15
N_DIAS	NUMBER	Yes	16
MEDIA_RECURSOS	NUMBER	Yes	17

JAEN_V12_DEM_AMP_LAB_AG_CTRL_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	4
TEMPERATURA_MAXIMA	NUMBER	Yes	5
TEMPERATURA_MINIMA	NUMBER	Yes	6
TEMPERATURA_MEDIA	NUMBER	Yes	7
HUMEDAD_MAXIMA	NUMBER	Yes	8
HUMEDAD_MINIMA	NUMBER	Yes	9

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
HUMEDAD_MEDIA	NUMBER	Yes	10
VELOCIDAD_VIENTO	NUMBER	Yes	11
RADIACION	NUMBER	Yes	12
PRECIPITACION	NUMBER	Yes	13
ETO	NUMBER	Yes	14
N_RECURSOS	NUMBER	Yes	15
N_DIAS	NUMBER	Yes	16
MEDIA_RECURSOS	NUMBER	Yes	17

JAEN_V12_DEM_AMP_LAB_AG_DTMN_YY

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
ANO	VARCHAR2(4 BYTE)	Yes	1
MES	VARCHAR2(2 BYTE)	Yes	2
SEMANA_MES	VARCHAR2(1 BYTE)	Yes	3
TIPO_DIA	VARCHAR2(9 BYTE)	Yes	4
TEMPERATURA_MAXIMA	NUMBER	Yes	5
TEMPERATURA_MINIMA	NUMBER	Yes	6
TEMPERATURA_MEDIA	NUMBER	Yes	7
HUMEDAD_MAXIMA	NUMBER	Yes	8
HUMEDAD_MINIMA	NUMBER	Yes	9
HUMEDAD_MEDIA	NUMBER	Yes	10
VELOCIDAD_VIENTO	NUMBER	Yes	11
RADIACION	NUMBER	Yes	12
PRECIPITACION	NUMBER	Yes	13
ETO	NUMBER	Yes	14

COLUMN_NAME	DATA_TYPE	NULLABLE	COLUMN_ID
N_RECURSOS	NUMBER	Yes	15
N_DIAS	NUMBER	Yes	16
MEDIA_RECURSOS	NUMBER	Yes	17