Diseño e implementación de un modelo colaborativo de información entre la población y los agentes de salud pública en el departamento del Atlántico

Design and implementation of a collaborative model of information between the population and public health agents in the department of Atlantic

K. Acuña*, D. Amador*, D. Mercado*, L. Vega*, Adonai Meza** J. Calabria***
{kevin.acuna, Daniel.amador , daniel.mercado, luis.vega1} @unisimon.edu.co – { } @unisimonbolivar.edu.co
*Estudiante de Ingeniería de sistemas **Estudiante de maestría e ingeniería de sistemas y computación ***Profesor investigador del grupo
Universidad Simón Bolívar, Barranquilla-Colombia.

Resumen La secretaría de salud en el departamento del Atlántico actualmente no cuenta con un modelo de captura de datos, que permita abstraer información de los ciudadanos referentes a antecedentes de enfermedades o conductas que influyan en la proliferación de una enfermedad no transmisible, es por esto por lo que es necesario fomentar la participación de las personas con los agentes de salud pública por medio de una plataforma por donde se diligencie esta información. Por tanto, se busca diseñar una aplicación web para la recolección de datos tomando como referencia el método progresivo de la OMS para la vigilancia de factores de riesgo de las enfermedades y el enfoque psicológico, para gestionar la información que será insumo para realizar un análisis detallado utilizando una inteligencia de negocios y extraer los factores de incidencia y su impacto real en el sistema de prevención de las enfermedades no transmisibles.

Palabras clave: *Enfermedades no Transmisibles, Enfermedades Crónicas, Prevención, Inteligencia de Negocios, Deep Learning.*

Abstract The Ministry of Health in the Atlantic Department does currently not have a data capture model, which allows abstract information from citizens concerning a history of diseases or behaviors that influence the proliferation of a non-communicable disease, this is why it is necessary to promote the participation of people with public health agents through a platform where this information is gathered. Therefore, it is necessary to design a web application for data collection by referring to the progressive method of WHO for the monitoring of risk factors for diseases and psychological approach, to manage the information that will be input to perform a detailed analysis Using a business intelligence and extracting the incidence factors and their actual impact on the prevention system of noncommunicable diseases.

Keywords: *Noncommunicable Diseases, Chronic Diseases, Prevention, Business Intelligence, Deep Learning.*

I. INTRODUCCIÓN

Actualmente muchas clínicas poseen herramientas informáticas desactualizadas las cuales generan algunos reportes, pero con el grave problema que crean métricas iguales pero con resultados diferentes lo cual afecta los informes de gestión dando una información que no es confiable ni precisa. En otros casos ni siquiera se cuenta con herramientas informáticas o incluso información digitalizada que permita que se generen informes de manera rápida y efectiva sobre el estado actual de los pacientes atendidos, sus principales patologías o información personal de los mismos. [1]

Las enfermedades no transmisibles o por sus siglas ENT matan según cifras de la Organización Panamericana de la Salud alrededor de 41 millones de personas cada año, y teniendo en cuenta cifras globales adquiridas por esta misma organización se puede afirmar que un 71% de las muertes tienen algún tipo de relación con las ENT, siendo las personas entre los 30 y 69 años los más afectados por esta enfermedad y los cuales representan el 85% de las muertes.

Asimismo, las ENT en el continente americano también tiene un serio impacto de tal forma que se llega a una cifra de 5,5 millones de muertes por año donde 2,2 millones de muertes corresponden a personas que tenían menos de 70 años, y que tienen como factor común una dificultad económica, ya que se tienden a presentar estas enfermedades en países de ingresos bajos o medios. Además, de la repercusión del factor económico en la calidad de vida y la aparición de ENT, se tienden a desarrollar unas enfermedades más que otras como son las cardiovasculares(17,9 millones de personas) y el cáncer(9,0 millones de personas), las cuales no necesariamente tiene relación con la economía del país, pero sí con sus hábitos de vida como seria el consumo de tabaco, alcohol, la inactividad física y las dietas no saludable, lo cual multiplica y pone en riesgo de contraer o morir por un ENT. [2]

En base a lo anterior, se justifica la necesidad de concientizar a la población sobre los riesgos que pueden traer ciertos hábitos y conductas como el consumo de cigarrillo, alcohol o vida sedentaria en el desarrollo de enfermedades crónicas, sin embargo, la secretaria de salud no cuenta actualmente con un modelo de captura de datos, donde se permita informar a las personas que tan probable es que se encuentren en riesgo de desarrollar una enfermedad crónica en base a sus antecedentes familiares y conductas. De aquí nace la necesidad de construir una metodología para la recolección de datos referente a los hábitos, conductas y cuadros clínicos, utilizando de base una muestra de la población del departamento del Atlántico de 60 ciudadanos, a partir de estos datos, se aplicará un modelo de hechos dimensionales y a partir de una inteligencia de negocios, se transformara estos datos crudos en conocimiento dirigido en saber cómo influyen estos registros en el aumento de riesgo del desarrollo de una enfermedad crónica.

II. ESTADO DEL ARTE

2.1. Inteligencia de negocios

Una definición para inteligencia de negocios o BI, por sus siglas en inglés, según el Data Waterhouse Institute, lo define como la combinación de tecnología, herramientas y procesos que me permiten transformar mis datos almacenados en información, esta información en conocimiento y este conocimiento dirigido a un plan o una estrategia comercial. La inteligencia de negocios debe ser parte de la estrategia empresarial, esta le permite optimizar la utilización de recursos, monitorear el cumplimiento de los objetivos de la empresa y la capacidad de tomar buenas decisiones para así obtener mejores resultados. [3]

2.2. Inteligencia de negocios en el sector de la salud

En el servicio de la salud, se pueden encontrar diversas metodologías y técnicas de las cuales se busca:

Disminuir el error médico y mejorar los procesos de salud es prioridad de todo el personal sanitario. En este contexto surgen los "Sistemas Clínicos de Soporte para la Toma de Decisiones" (CDSS), los cuales son un componente fundamental en la informatización de la capa clínica. Con la evolución de las tecnologías gran cantidad de datos han podido ser estudiados y clasificados a partir de la minería de datos. Una de las principales ventajas de la utilización de esta, en los CDSS, ha sido su capacidad de generar nuevos conocimientos. Con este fin se propone, mediante la combinación de dos modelos matemáticos, cómo se puede contribuir al diagnóstico de enfermedades usando técnicas de minería de datos. Para mostrar los modelos utilizados se tomó como caso de estudio la hipertensión arterial. El desarrollo de la investigación se rige por la metodología más utilizada actualmente en los procesos de Descubrimiento de Conocimiento en Bases de Datos: CRISP-DM 1.0, y se apoya en la herramienta de libre distribución WEKA 3.6.2, de gran prestigio entre las utilizadas para el modelado de minería de datos. Como resultados se obtuvieron diversos patrones de comportamiento con relación a los factores de riesgo a sufrir hipertensión mediante técnicas de minería de datos. [4]

La medición y evaluación del estado de salud en las comunidades de todo el mundo es un desafío masivo de tecnología de la información. La Evaluación Integral para el Seguimiento de la Salud Comunitaria (CATCH) proporciona métodos sistemáticos para la evaluación a nivel comunitario que es invaluable para la asignación de recursos y la formulación de políticas de atención médica. CATCH se basa en indicadores de estado de salud de múltiples fuentes de datos, utilizando un marco comparativo innovador y un proceso de evaluación ponderado para producir una lista ordenada por orden de importancia de desafíos críticos de atención de salud comunitaria. El enfoque a nivel de la comunidad tiene el propósito de capacitar a los responsables locales de la toma de decisiones proporcionando una metodología clara para organizar e interpretar los datos relevantes sobre la atención médica. La extensa experiencia de campo con los métodos CATCH, en combinación con la experiencia en tecnología de almacenamiento de datos, ha llevado a una aplicación innovadora

de la tecnología de la información en el ámbito de la atención de la salud. El almacén de datos permite producir un conjunto básico de informes a un costo razonable para el uso comunitario. Además, la funcionalidad de procesamiento analítico en línea (OLAP) se puede utilizar para obtener una comprensión más profunda de problemas específicos de atención de la salud. El almacén de datos junto con los métodos de difusión habilitados para la Web permite que la información se presente en una variedad de formatos y se distribuye más ampliamente en la comunidad de toma de decisiones. En este artículo, nos centramos en los desafíos técnicos de diseñar e implementar un almacén de datos eficaz para la información de atención de salud. A lo largo de la discusión se utilizan ilustraciones de los diseños de datos reales y los formatos de informes del almacén de datos CATCH. Conclusiones de las investigaciones en curso en el almacenamiento de datos de salud y la toma de decisiones de salud comunitaria. [5]

La medicina basada en la evidencia es una nueva dirección en la salud moderna. Su tarea es prevenir, diagnosticar y medicar enfermedades usando evidencia médica. Los datos médicos sobre una gran población de pacientes se analizan para realizar la gestión sanitaria y la investigación médica. Con el fin de obtener la mejor evidencia para una enfermedad dada, la experiencia clínica externa, así como la experiencia clínica interna para estar disponibles para los profesionales de la salud en el momento adecuado y de la manera correcta. El conocimiento externo basado en la evidencia no puede aplicarse directamente al paciente sin ajustarlo al estado de salud del paciente. Basamos un enfoque basado en almacenes de datos como una solución adecuada para la integración de fuentes de datos externas basadas en evidencia en el sistema de información clínica existente y técnicas de minería de datos para encontrar terapia apropiada para un paciente y una enfermedad dados. [6]

2.3. Metodología para una bodega de datos

La metodología de construcción de bodegas de datos propuesta se centra en una estructura tecnológica diseñada para que la organización disponga, en forma integrada y estandarizada, de la información correspondiente a la operación de la empresa, así como, proporciona a los usuarios que tienen a su cargo la toma de decisiones, las herramientas adecuadas, para que a través de consultas rápidas ellos mismos accedan a la información requerida. [7]

Para su adecuada estructuración es necesario reconocer un conjunto de operaciones y convertirlas en un sistema orientado al sujeto donde el instante de tiempo en que ocurre la acción es de importancia vital para su análisis, el enfoque inicial se dirige hacia el conocimiento de los distintos orígenes de datos (motores administradores de base de datos, sistemas de archivos etc.) de los cuales serán extraídos los datos que en nuestra bodega se convertirán en información [8]

2.4. Modelo de hechos dimensionales estrella

El esquema de estrella es un enfoque de modelado maduro ampliamente adoptado por los almacenes de datos relacionales. Requiere que los modeladores clasifiquen las tablas del modelo como dimensiones o hechos. [9]

Las tablas de dimensiones describen entidades empresariales (las cosas que se modelan). Las entidades pueden incluir productos, personas, lugares y conceptos, incluido el propio tiempo. La tabla más coherente de un esquema de estrella es una tabla de dimensiones de fecha. Además, contiene una columna (o columnas) de clave que actúa como identificador único y columnas descriptivas. Las tablas de hechos pueden almacenar observaciones o eventos, y pueden ser pedidos de ventas, existencias, tasas de cambio, temperaturas, etc. Una tabla de hechos contiene columnas de clave de dimensiones relacionadas con las tablas de dimensiones y columnas de medida numéricas. Las columnas de clave de dimensiones determinan la dimensionalidad de una tabla de hechos, mientras que los valores de clave de dimensiones determinan la granularidad de una tabla de hechos. [9]

2.5. Herramientas para la ETL

Power BI es una herramienta permite analizar e interactuar con una cantidad masiva de datos dentro de Excel, recabando información a través de lenguaje natural y peticiones del tipo pregunta-respuesta. Además, tiene la capacidad de buscar fácilmente, transformar y tener acceso a datos dentro y fuera de la organización, desde cualquier lugar y en casi cualquier dispositivo. Asimismo, podrás analizar y crear atractivas visualizaciones interactivas. Tiene la capacidad de potenciar características de Excel mediante: Power Query, Power Map, Power Pivot, y Power View, los cuales son herramientas que permiten tener un mayor agrado visual al leer diagramas. [10]

Qlik sense es una aplicación lanzada por QlikTech, que se utiliza específicamente para visualizar y analizar los datos. Ayuda a crear paneles e informes interactivos, y también a extraer los datos de varias fuentes de datos. En su mayoría, sirve como una plataforma de análisis para desarrolladores técnicos y no técnicos y funciona mejor para la visualización de datos, ya que tiene gráficos aumentados que ayudan en la visualización eficaz de los datos. [11]

2.6. Aprendizaje profundo

El deep learning es un tipo de machine learning que entrena a una computadora para que realice tareas como las hacemos los seres humanos, como el reconocimiento del habla, la identificación de imágenes o hacer predicciones. En lugar de organizar datos para que se ejecuten a través de ecuaciones predefinidas, el deep learning configura parámetros básicos acerca de los datos y entrena a la computadora para que aprenda por cuenta propia reconociendo patrones mediante el uso de muchas capas de procesamiento. [12]

En el año 2010, gracias al enorme avance de la capacidad de computación de los ordenadores, se populariza el aprendizaje profundo o deep learning, un concepto que ya había aparecido en el año 1974 de la mano de Paul Werbos. Esto es una rama del aprendizaje automático, machine learning, que se basa en el uso de redes neuronales artificiales, artificial neural networks o ANN, que intentan imitar el funcionamiento de nuestro cerebro. Los algoritmos de aprendizaje profundo no precisan disponer de reglas programadas previamente, sino que aprenden por sí mismos qué características son las mejores para efectuar una tarea. [13]

Big data es un término que describe el gran volumen de datos – estructurados y no estructurados – que inundan una empresa todos los días. Pero no es la cantidad de datos lo importante. Lo que importa es lo que las organizaciones hacen con los datos. El big data puede ser analizado para obtener insights que conlleven a mejores decisiones y acciones de negocios estratégicas. [14]

Para entender qué significa realmente "big data", resulta útil conocer ciertos antecedentes históricos. A continuación, ofrecemos la definición de Gartner, de aproximadamente 2001 (y que continúa siendo la definición de referencia): Big data son datos que contienen una mayor variedad y que se presentan en volúmenes crecientes y a una velocidad superior. Esto se conoce como "las tres V".

Dicho de otro modo, el big data está formado por conjuntos de datos de mayor tamaño y más complejos, especialmente procedentes de nuevas fuentes de datos. Estos conjuntos de datos son tan voluminosos que el software de procesamiento de datos convencional sencillamente no puede administrarlos. Sin embargo, estos volúmenes masivos de datos pueden utilizarse para abordar problemas empresariales que antes no hubiera sido posible solucionar. [15]

2.7. Enfermedades no transmisibles

La Organización Mundial de la Salud (OMS), define las enfermedades no transmisibles (ENT) o crónicas, como “afecciones de larga duración con una progresión generalmente lenta. Entre ellas destacan: las enfermedades cardiovasculares (por ejemplo, los infartos de miocardio o accidentes cerebrovasculares); el cáncer; las enfermedades respiratorias crónicas (por ejemplo, la neumopatía obstructiva crónica o el asma); y la diabetes” [16]

Existen pruebas científicas concluyentes de que una dieta saludable y una actividad física adecuada (es decir, alrededor de 30 minutos de actividad física de intensidad moderada unos cinco días a la semana) desempeñan un papel importante en la prevención de estas enfermedades (p.41). [17]

De las estrategia que la 57ª Asamblea Mundial de la Salud OMS aprueban a nivel mundial destaco la que sería base de esta investigación que es la del tópico de Educación, comunicación y concientización del público, afirma “Una buena base para la adopción de medidas es la difusión y comprensión de conocimientos adecuados sobre la relación entre el régimen alimentario, la actividad física y la salud, el aporte y el gasto energético, así como sobre decisiones saludables en materia de productos alimenticios. Los expertos gubernamentales, las organizaciones no gubernamentales y comunitarias y las industrias apropiadas deben formular y transmitir mensajes consecuentes, coherentes, sencillos y claros.

Dichos mensajes han de comunicarse por diversos canales y de una forma que responda a las características de la cultura local, así como a la edad y el sexo de las personas. Es posible influir concretamente en los comportamientos desde las escuelas, los entornos laborales y las instituciones docentes y religiosas, así como por conducto de las organizaciones no gubernamentales, los líderes comunitarios y los medios de difusión.

Los Estados Miembros deben formar alianzas para dar amplia difusión a mensajes apropiados y eficaces sobre el régimen alimentario y la actividad física. La educación en materia de alimentación y actividad física, así como la adquisición de conocimientos básicos sobre los medios de comunicación desde la escuela primaria, son importantes para promover dietas más sanas y resistir a las modas alimentarias y a la información engañosa sobre

estas cuestiones. También se ha de prestar apoyo a la adopción de medidas encaminadas a mejorarlos conocimientos básicos en materia de salud, teniendo en cuenta las circunstancias culturales socioeconómicas locales. Es preciso evaluar periódicamente las campañas de comunicación." [18]

Por lo anterior se parte de esa teoría y se hace necesario definir con claridad la construcción y entendimiento conceptual del objeto de investigación en la población objetivo; medir realmente su impacto en la prevención de enfermedades no transmisible; entonces así podremos aportar a ese postulado teórico con una definición más apropiada en nuestro entorno o talvez encontraremos otros factores relevantes que aporte a una eficacia y eficiencia en el sistema de prevención del mismo.

III. METODOLOGÍA DE LA INVESTIGACIÓN

La implementación de este modelo colaborativo se fundamenta en los datos recolectados al tomar una muestra de las personas que habitan en el departamento del atlántico, el proceso para la misma abarca cuatro importantes etapas, son identificadas como:

3.1. Análisis:

Para alcanzar los objetivos de la investigación se planteó la realización de socialización y aplicación del formulario de enfermedades no transmisibles (ENT) el cual en teoría consta de 60 preguntas de las cuales 50 son de opción múltiple con única respuesta en la mayoría de su composición, y 10 preguntas con respuesta argumentativa, esto con el fin de recolectar información acerca de los antecedentes de cáncer, epoc, hta, diabetes, y datos relacionados con el manejo de la salud de los habitantes de los municipios, corregimientos u otro ente territorial adscrito al Departamento del Atlántico.

3.2. Implementación de métodos de investigación apropiados:

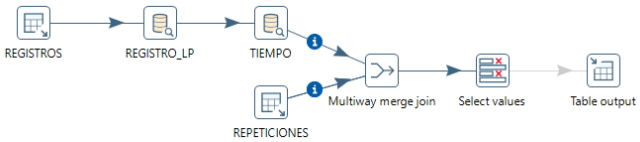
Al haber analizado literaturas disponibles, se desarrollara y ultimará los métodos de investigación apropiados en colaboración con las habitantes y agentes de salud pública del departamento y locales, generando espacios colaborativos para gestionar información que será insumo para realizar un análisis detallado y en conjunto de los agentes de salud pública validar los resultados preliminares, identificando los factores de incidencia y su impacto real en el sistema de prevención de las enfermedades no transmisibles constituyendo así a la construcción de un panorama de prevención que será monitoreado a través de indicadores aprobados por la secretaria de salud departamental para posteriormente sensibilizar mediante la realización de reuniones a nivel de local y departamental, garantizando una retroalimentación al sistema de prevención de enfermedades no transmisibles del Departamento del Atlántico.

3.3. Análisis y recolección de datos

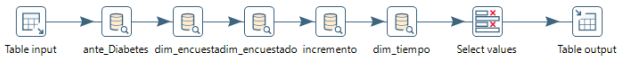
En la implementación de este, los datos suministrados por los usuarios se procesaron mediante un proceso de ETL(Extracción, Transformación y Carga) el cual consta de extraer al información almacenada en la base de datos relacional para posteriormente mediante el aplicativo seleccionado para esta tarea Pentaho, se hiciera el proceso de extracción de datos al propio aplicativo, consecutivamente las transformaciones requeridas y por último el cargue a la bodega de datos. Asimismo, en secuencia a lo anterior se manejó la herramienta Qlink Sense como método para la generación de un dashboard el cual se utilizó para generar estadísticas en base a filtros aplicados, dando como resultado elementos para la toma de decisiones como son los KPI para la muestra de datos, diagrama de caja, y linear, además de un mapa en el cual se ven reflejado la incidencia de los valores aplicados en el filtro.

Estructura de las dimensiones creadas en Pentaho

Estructura fact table antecedentes_cancer



Estructura fact table ciudadano



Estructura fact table antecedente_diabetes

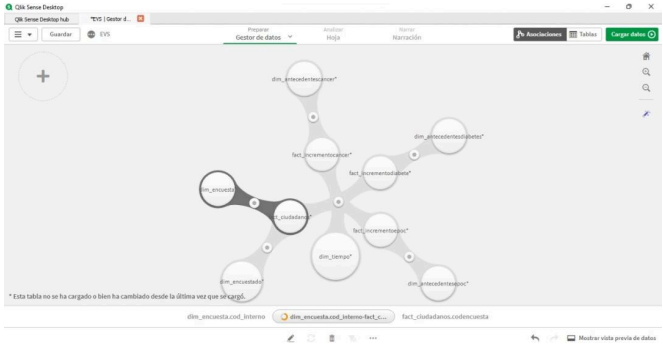
Estructura fact table antecedente_epoc

3.4. Publicación de resultados:

Al final se realizaría la difusión y publicación de los factores de incidencia y su impacto en la prevención de enfermedades no transmisible en la población vulnerable del departamento del Atlántico.

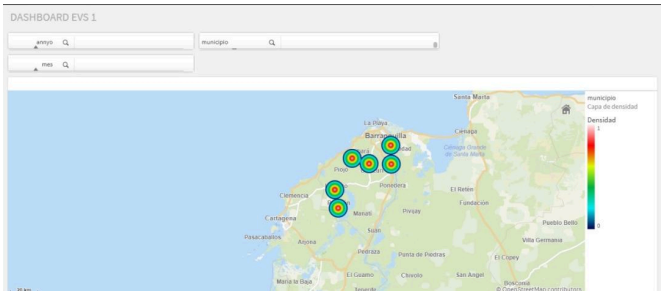
IV. RESULTADOS OBTENIDOS

En base a los datos almacenados en la bodega de datos se logró la creación del dashboard que contiene información relevante para la toma de decisiones y posibles medidas preventivas en relación con las ENT



Ventana de creación de asociaciones para Qlink Sense, en este apartado se crean las asociaciones bajo la cual se trabaja en el aplicativo y mediante esta se realiza el cargue de datos los cuales manejaremos a posteriori.

Dashboard datos de ENT, este dashboard cuenta con graficas de línea y de cajas, además de con elementos KPI los cuales nos indican las incidencias de las ENT de acuerdo con los filtros seleccionados.



Dashboard mapa incidencias de ENT, en base a la información suministrada por el mapa de incidencias se puede apreciar la presencia de cada ENT de acuerdo con el filtro seleccionado en relación con el tiempo (igualmente seleccionado en el filtro o por el total desde el inicio hasta el fin de la toma de datos), con relación al mapa se muestra el origen de las encuestas realizadas en el Departamento del Atlántico.

V. CONCLUSIONES

- Para la recolección de información se recomienda utilizar alguna herramienta online o la utilización de un framework de desarrollo backend para acelerar el proceso del aplicativo web, donde se realizarían las encuestas, ya que para la fase de extracción, transformación y carga de datos en la bodega, se contaban con menos de 100 encuestas realizadas, lo que disminuyó la precisión de los resultados obtenidos en la herramienta de inteligencia de negocios.

- El uso de la herramienta de inteligencia de negocios Qlik sense, fue fundamental para la creación de un dashboard donde se pudo observar de manera dinámica el comportamiento y de los datos de muestra y su distribución en el departamento del Atlántico.

- Al usar una herramienta para la inteligencia de negocios se permite la escalabilidad del proyecto, ya que después de realizados los procesos de transformación, extracción y carga de la base de datos, se puede importar fácilmente la bodega sin tener que reajustar nada de los dashboard en la herramienta.

REFERENCIAS

[1] L. Ardila, J. Angulo, B. Rangel y R. Samir, Diseño y desarrollo de un dashboard utilizando una bodega de datos integrada con una aplicación web para realizar consultas especializadas y elaboración de reportes inteligentes sobre la gestión de pacientes de las clínicas..

[2] OMS, Enfermedades no transmisibles, 2021.

[3] Oracle, ¿Qué es Inteligencia de Negocios?.

[4] E. A. O. Carrascal, A. I. O. Carrascal y G. L. V. Saldarriaga, «MINERÍA DE DATOS: APORTES Y TENDENCIAS EN ELSERVICIO DE SALUD DE CIUDADES INTELIGENTES,» 2015.

[5] F. D. Hernández y Y. S. Corales, «Técnicas de minería de datos aplicadas al diagnóstico de entidades clínicas,» *Reviste Cubana de Infórmatica Médica,* pp. Vol. 4 - no. 2, 2012.

[6] N. Stolba y M. Tjoa, «The relevance Data Warehousing and Data,» World academy of science, Engineering and technology international journal of healt and medical engineering, 2007.

[7] Witten Ian y Frank Eibe, Data Mining: Practical Machine Learning Tools and Techniques, 2005.

[8] J. C. C. Sarmiento, Construccion y poblamiento de un datawarehouse basado en el paradigma de base de datos objeto relacional, 2011.

[9] P. Myers y O. , Descripción de un esquema de estrella e importancia para Power BI, Microsoft, 2019.

[10] Softeng, «La nueva herramienta de Office 365 para trabajar con tus datos: Power BI,» 4 Octubre 2013. [En línea]. Available: https://www.softeng.es/es-es/blog/power-bi-la-nueva-herramienta-de-office-365-para-trabajar-con-datos-masivos.html.

[11] M. Vinod, «Qlik Sense tutorial,» MindMajix, 23 agosto 2021. [En línea]. Available: https://mindmajix.com/qlik-sense-tutorial. [Último acceso: 11 11 2021].

[12] SAS, «SAS,» [En línea]. Available: https://www.sas.com/es_mx/insights/analytics/deep-learning.html.

[13] I. Menéndez, «Babel,» Aplicando deep learning en el ámbito de la salud, 16 3 2021. [En línea].

[14] SAS, «SAS,» [En línea]. Available: https://www.sas.com/es_mx/insights/big-data/what-is-big-data.html.

[15] ORACLE, «ORACLE,» [En línea]. Available: https://www.oracle.com/mx/big-data/what-is-big-data/.

[16] OMS, «Enfermedades no transmisibles,» [En línea]. Available: https://www.who.int/topics/noncommunicable_diseases/es/.

[17] OMS, «Prevención de las enfermedades no transmisibles en el lugar de trabajo a través del regimen alimentario y la actividad física,» Foro Económico Mundial, 2008.

[18] OMS, «Estrategia mundial sobre régimen alimentario, actividad física y salud,» 2004. [En línea]. Available: https://www.who.int/dietphysicalactivity/strategy/eb11344/strategy_spanish_web.pdf.