

PROPUESTA DE UNA APLICACIÓN PARA REALIZAR LA PREDICCIÓN DE LA APARICIÓN DE PADECIMIENTOS INCAPACITANTES EN PACIENTES CON DIABETES TIPO II

*Sahir Omar Burciaga Sosa¹, Beatriz A. Olivares Zepahua², Giner Alor Hernandez³, Ana María
Chávez Trejo⁴*

1,2,3,4 Instituto Tecnológico de Orizaba, Veracruz, México

*1sburciaga@acm.org, 2bolivares@ito-depi.edu.mx, 3galor@ito-depi.edu.mx, 4achavez@ito-
depi.edu.mx*

Abstract

La diabetes representa un grave problema de salud en México; constituye la segunda causa de mortalidad en el país. El padecimiento de esta enfermedad aumenta el riesgo de desarrollar diversas complicaciones como problemas renales, ceguera, daño a los nervios y enfermedades cardíacas, las cuales afectan la calidad de vida del paciente y, en el peor de los casos, conducen a la muerte. Para solventar tal problemática el presente trabajo propone el desarrollo de una aplicación que, utilizando Minería de Datos, realice la predicción de la aparición de padecimientos incapacitantes en pacientes con Diabetes Mellitus Tipo II en la región de Orizaba así como la generación de una ontología que provea una lista de cuidados preventivos para retrasar/evitar la aparición de esos padecimientos. Con el desarrollo de esta aplicación se mejoraría la identificación temprana de padecimientos crónicos en los pacientes con Diabetes Mellitus Tipo II y se diseñarían apropiadas intervenciones por parte de los profesionales de la salud.

PALABRAS CLAVE: Minería de Datos, diabetes, ontología, clasificación, predicción, WEKA.

Introducción

La diabetes mellitus, o simplemente diabetes, es una enfermedad que representa un grave problema de salud en México y en el mundo. La Encuesta Nacional de Salud y Nutrición (ENSANUT) en el 2012 reveló que el 9% de la población mexicana adulta padece esta enfermedad [1] y se estima que esta cantidad probablemente sea del doble debido a que muchas personas aún no son diagnosticadas. Por otro lado, el Instituto Nacional de Estadística, Geografía e Informática (INEGI), en el censo llevado a cabo en 2010, reveló que la diabetes es responsable del 14% de las muertes en el país representando la segunda causa de mortalidad en México [2].

Además, la diabetes aumenta el riesgo de desarrollar diversas complicaciones como problemas renales, ceguera, daño a los nervios, daño a los vasos sanguíneos y enfermedades cardíacas,

las cuales afectan la calidad de vida del paciente y, en el peor de los casos, conducen a la muerte. La diabetes representa la principal causa de ceguera en México y una de las principales causas de insuficiencia renal [3].

Los organismos de salud pública buscan constantemente la implementación de Tecnologías de Información y Comunicaciones (TICs) para facilitar la gestión, el seguimiento y el diagnóstico temprano de pacientes; entre esas tecnologías se encuentran técnicas de Minería de Datos que se utilizan como apoyo al médico experto para diagnosticar enfermedades y el uso de ontologías para ofrecer alternativas de tratamiento y de cuidados sanitarios a los pacientes. Por su lado, la Minería de Datos permite analizar una gran cantidad de información sobre un conjunto de datos que de otra forma sería difícil de analizar, en este caso por los profesionales de la salud, y con ello entrenar un modelo con el cual posteriormente es posible realizar predicciones respecto a la aparición de otros padecimientos. Por otro lado, las ontologías permiten representar conocimiento dentro de un dominio en particular, por ejemplo dentro del ámbito de una enfermedad, y proporcionar información correspondiente a ella.

Conceptos

A continuación se describen los conceptos básicos necesarios para un mejor entendimiento del presente documento.

Diabetes mellitus tipo II y enfermedades relacionadas

La Diabetes Mellitus (DM) es un conjunto de enfermedades que da como resultado defectos en la secreción de insulina. Existen dos tipos de DM: Tipo I que es la incapacidad del cuerpo para producir insulina y Tipo II que es una condición en la que el cuerpo no usa adecuadamente la insulina que produce. Esta última es la más común en todo el mundo, representa entre el 90 y el 95% de la población diabética [4] [5] [6]. Las personas que padecen DM tienen una alta probabilidad de presentar diversas complicaciones, a continuación se mencionan algunas de las más comunes: El pie diabético es el trastorno ulcerativo de los pies que se produce por el daño en nervios y arterias del pie ocasionados por la hiperglucemia sostenida. La retinopatía diabética es una complicación ocular de la diabetes que se origina por el deterioro de los vasos sanguíneos que irrigan la retina del fondo del ojo. Es una de las causas más frecuente de ceguera entre los adultos [4]. La nefropatía diabética se caracteriza por generar daños al sistema glomerular, ocasionando un crecimiento del mismo y produciendo un aumento en el tamaño de los riñones. Es una de las causas de falla renal y de la necesidad de que un paciente reciba hemodiálisis. Las complicaciones renales representan en México la primera causa de hospitalización en pacientes diabéticos [7].

Minería de Datos

Mehmed Kantardzic define a la Minería de Datos (MD) como “el proceso para encontrar información nueva, valiosa y no trivial en grandes volúmenes de datos” [8]. En otras palabras, la MD permite extraer conocimiento útil y desconocido que se encuentra oculto en los datos. La MD se aplica en una gran variedad de tipos de repositorios como son base de datos relacionales, almacenes de datos, archivos planos, entre otros [9]. Las tareas principales de la MD son la predicción y la descripción [9].

Clasificación y predicción

La clasificación y la predicción permiten analizar los datos y extraer modelos que ayudarán a predecir el comportamiento de estos. La principal diferencia entre estos conceptos radica en que la clasificación se utiliza cuando se desea obtener valores discretos y la predicción cuando se desean conocer valores continuos [10]; sin embargo algunos autores, e incluso herramientas como las de Microsoft, usan solamente el término predicción sin importar el tipo específico de valor.

Ontología

Una de las definiciones de ontologías más aceptada es la propuesta por Gruber en 1993 y posteriormente mejorada por Studer en 1998 definiéndola de la siguiente manera “una ontología es una especificación formal y explícita de una conceptualización compartida”. Donde Especificación formal se refiere a que la ontología debe ser legible para la computadora. Conceptualización es un modelo, una abstracción del mundo que se desea representar para algún propósito. Explícita se refiere a que el modelo debe de estar bien definido, para ello se hace uso de relaciones, funciones, restricciones, axiomas y conceptos. Y finalmente compartida ya que captura conocimiento aceptado por un grupo. Una ontología representa conocimiento dentro de un dominio en particular mediante un conjunto de conceptos y sus relaciones. [11].

Ontología para los Trastornos del Metabolismo de la Glucosa

El Centro Nacional de Ontologías Biomédicas (National Center for Biomedical Ontology, NCBO) es un consorcio internacional de investigadores comprometidos a mejorar la salud. Por medio de su portal permiten el acceso al conocimiento y datos biomédicos utilizando principalmente ontologías que sirven de apoyo para la ciencia biomédica y atención clínica [12]. La ontología para los trastornos de metabolismo de la glucosa (Ontology of Glucose Metabolism Disorder, OGMD) pertenece al NCBO y se desarrolló para apoyar el estudio trasnacional sobre diabetes mellitus [13].

Propuesta

Para intentar solventar el problema de salud que genera la diabetes y sus complicaciones en México, el presente trabajo propone el desarrollo de una aplicación que, utilizando Minería de Datos, realice la clasificación (predicción) de la aparición de padecimientos incapacitantes en pacientes con Diabetes Mellitus Tipo II en la región de Orizaba así como la generación de una ontología que provea una lista de cuidados preventivos para retrasar/evitar la aparición de esos padecimientos. En la figura 1.1 se muestra, de una forma muy general, el proceso para la elaboración de la aplicación y las tecnologías que son necesarias para su desarrollo. Como primer punto se observa que la construcción de la aplicación se encuentra dividida en tres módulos. El primero representa aquellas tecnologías y actividades que conciernen a tareas propias de Minería de Datos. El segundo engloba a aquellas tareas que permiten la creación y manejo de ontologías. Y finalmente, el tercero representa la aplicación Web con la cual se podrá conjuntar a los primeros dos módulos para realizar la predicción de padecimientos incapacitantes junto con su lista de cuidados preventivos.

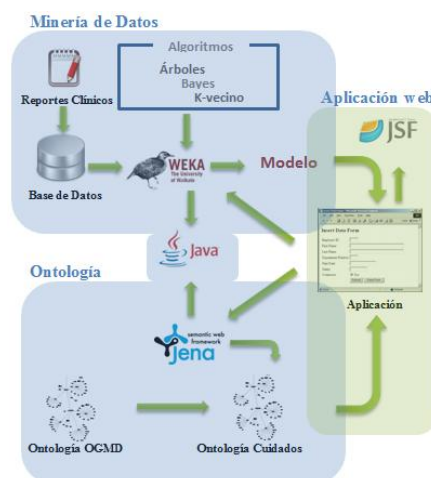


Figura 1 Ilustración del proceso para la elaboración de la aplicación propuesta.

Primeramente, para la elaboración del módulo de Minería de Datos, se recolectan todos los reportes clínicos disponibles en la institución médica donde se realice el estudio y con la ayuda de un profesional de la salud se identifican aquellas variables que se consideren importantes para la predicción de padecimientos incapacitantes en pacientes con diabetes mellitus tipo II con el fin de almacenarlas en la base de datos para su posterior procesamiento. Dicho procesamiento consiste en la detección y tratamiento de valores atípicos, faltantes o perdidos dependiendo de las condiciones en que se encuentren los datos extraídos de los reportes clínicos y la eliminación de información errónea o irrelevante. Al momento de redactar el artículo no se tiene un convenio con una institución de salud, éste se hará posteriormente en

una institución de salud de la región de Orizaba. Posteriormente, con el apoyo de la herramienta WEKA, que es un software que permite realizar diversas tareas de minería de datos como clasificación (predicción), se procederá a la elaboración de un modelo utilizando diferentes algoritmos de Minería de Datos como Árboles de Decisión, Naive Bayes y Vecino más Cercano entre otros, con la finalidad de comparar sus resultados y obtener el modelo que presente un mejor desempeño, con base en su nivel de certeza, sensibilidad y especificidad. Finalmente este modelo será el que se utilice en la aplicación Web para realizar la predicción de la aparición de padecimientos incapacitantes.

Por otro lado, en el módulo de ontologías, también con la ayuda de un médico experto, se determinan aquellos conceptos que se consideren relevantes dentro del dominio de enfermedades incapacitantes y cuidados preventivos tomando en cuenta la ontología OGMD, descrita en el apartado anterior, y se procede a la creación de una ontología específica que contiene información detallada acerca de la diabetes y sus complicaciones. Para manejar las ontologías se hace uso de Jena que es un software para la gestión y consulta de ontologías. Finalmente con Jena se accederá a la ontología de cuidados preventivos que propone el tratamiento adecuado para un determinado padecimiento.

Por último, se contará con una aplicación Web, bajo la tecnología Java Server Faces (JSF), a la cual se integrarán el modelo de minería de datos y la ontología de cuidados preventivos. Esta aplicación permitirá introducir los valores necesarios para realizar la predicción y mostrar, en caso de predecir la aparición de una enfermedad incapacitante, una lista detallada de cuidados preventivos para dicha enfermedad.

Por otra parte es cierto que existen diversos artículos en México que abordan la problemática de la diabetes, pero ellos se centran principalmente en el análisis estadístico de la presencia de la enfermedad [14] [15] y las complicaciones que de ella derivan, como son la gestión de tratamientos de largo plazo, la necesidad de unidades médicas especializadas, los costos de hospitalización y el aumento de pensiones por incapacidad [16]. En otros artículos, a nivel internacional, se observa que la mayoría se basan en el diagnóstico de alguna enfermedad, en este caso la diabetes, pero muy pocos se centran en la aparición de alguna complicación en personas que ya padecen esta enfermedad. Y en la mayoría de los casos, una vez diagnosticada, no se ofrece un tratamiento o tipo de cuidado hacia el paciente.

Conclusiones

Es importante valorar el impacto de las Tecnologías de Información y Comunicación en el sector de la salud. El uso de técnicas de minería de y ontologías constituyen una solución viable para lograr combatir el grave problema de salud que representa la diabetes en México e intentar reducir la tasa de mortalidad, costos económicos y complicaciones que de ella se derivan.

Asimismo, después de establecer un proceso general para la elaboración de la aplicación a través de la identificación de los módulos de minería de datos, de ontologías y de la aplicación web se fortalecen las bases para el posterior desarrollo de la aplicación propuesta.

Trabajo a futuro

Como trabajo a futuro se realizará el desarrollo de la aplicación. Se construirá el módulo de Minería de Datos, el módulo de ontologías y finalmente la aplicación Web utilizando la tecnología de JSF con la cual se podrá realizar la predicción de la aparición de algún padecimiento incapacitante junto con su tratamiento correspondiente.

Agradecimientos

Al Consejo Nacional de Ciencia y Tecnología, CONACyT, por el apoyo económico otorgado para la realización de la investigación.

Referencias

- [1] Instituto Nacional de Estadística y Geografía. "Banco de información". October 2013. [Online]. Available: <http://www3.inegi.org.mx/sistemas/biinegi/default.aspx>
- [2] ENSANUT "Encuesta Nacional de Salud y Nutrición 2012". October 2013. [Online]. Available: <http://ensanut.insp.mx/>
- [3] Secretaría de Salud. (2007). Programa de Acción Específico 2007 – 2012. Unidades de Especialidades Médicas UNEMES.
- [4] Organización Panamericana de la Salud / Panamá, "Guía para la atención integral de las personas con diabetes mellitus", Panamá : Organización Panamericana de la Salud, 2009.
- [5] K. Rajesh and V. Sangeetha, "Application of Data Mining Methods and Techniques for Diabetes Diagnosis", International Journal of Engineering and Innovative Technology (IJEIT), vol. 2, no. 3, 2012, ISSN: 2277-3754

- [6] Marcano-Cedeno, Alexis, and Diego Andina, "Data mining for the diagnosis of type 2 diabetes", In World Automation Congress (WAC), pp. 1-6. IEEE, June 2012, ISSN 2154-4824
- [7] J. Escobedo de la Peña and B. Rico Verdín, "Incidencia y letalidad de las complicaciones agudas y crónicas de la diabetes mellitus en México; Incidence and fatlity rates of short and long-term diabetes mellitus complications in Mexico", Salud pública Méx, vol. 38, no. 4, pp. 236-42,1996.
- [8] Mehmed Kantardzic. 2011. Data Mining: Concepts, Models, Methods, and Algorithms (2nd ed.). Wiley-IEEE Press.
- [9] Jiawei Han, Micheline Kamber, and Jian Pei. 2011. Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- [10] The University of Waikato. "Weka". September 2013. [Online]. Available: <http://www.cs.waikato.ac.nz/ml/weka/>
- [11] Staab, Steffen, and Rudi Studer. 2009. Handbook on ontologies. Springer.
- [12] National Center for Biomedical Ontology. About NCBO. Retrieved from <http://www.bioontology.org>
- [13] Lin, Yu. "Refining Ontology for Glucose Metabolism Disorders." ICBO. 2011.
- [14] Gustavo Olaiz Fernández, Rosalba Rojas, Carlos A Aguilar-Salinas, Juan Rauda and Salvador Villalpando, "Diabetes mellitus en adultos mexicanos. Resultados de la Encuesta Nacional de Salud 2000", Salud Publica Mex, vol. 49, no. s3, 2007
- [15] Escobedo de la Peña, Jorge and Beatriz Rico Verdín, "Incidencia y letalidad de las complicaciones agudas y crónicas de la diabetes mellitus en México; Incidence and fatlity rates of short and long-term diabetes mellitus complications in Mexico", Salud pública Méx, vol. 8 , no. 4, pp. 236-42, 1996.
- [16] Armando Arredondo and Esteban De Icaza , "Costos de la diabetes en América Latina: evidencias del caso Mexicano", Value Health, vol. 14, no.5 Suppl 1, pp. S85-8, 2011.

