TRABAJO ORIGINAL

MINERÍA DE DATOS EN DIABETES

DATA MINING IN DIABETES

Franklin Ábalos[1], José Retamosa[2], Mónica Roldán[3], Alejandro Ábalos[4], Patricia Lettari[5], Ruth Rojo[6]

RESUMEN

La minería de datos (MD) es una herramienta relativamente novedosa que permite descubrir relaciones o conexiones desconocidas en bases de datos. Su aplicación en medicina es escasa[1], y por ende en diabetes. Describimos sus nociones básicas, acompañadas de hipotéticos casos relacionados con diabetes, para aplicar en nuestra práctica diaria.

Palabras clave: minería de datos, diabetes, tecnología.

Revista de la Sociedad Argentina de Diabetes 2015; Vol. 49 (77-84)

ABSTRACT

Data mining is a relatively new tool that allows finding out unknown associations in databases. Its use in medicine is scarce and therefore in diabetes. We describe basic notions with hypothetic examples regarding diabetes, to be applied in our daily practice.

Key words: *data mining, diabetes, technology.*

Revista de la Sociedad Argentina de Diabetes 2015; Vol. 49 (77-84)

[1] Médico Especialista en Medicina Interna. Director Centro Municipal de Diabetes, Santiago del Estero (capital)
[2] Médico Especialista en Diabetes. Médico del Departamento de Diabetes del Hospital Regional, Santiago del Estero (capital)
[3] Médica Especialista en Clínica Médica. Médica del Departamento de Diabetes del Hospital Independencia, Santiago del Estero (capital)
[4] Médico Especialista en Nefrología. Médico del Departamento de Nefrología del Hospital Regional, Santiago del Estero (capital)
[5] Médica Especialista en Clínica Médica. Médica del Departamento de Clínica Médica del Hospital Independencia, Santiago del Estero (capital)
[6] Lic. en Nutrición. Nutricionista del Centro Calidad de Vida (Diabetes). Obra Social del Empleado Público Provincial, Santiago del Estero

Contacto del autor: Franklin Ábalos
E-mail: franklinabalos@gmail.com
Correspondencia: La Plata 456 (CP 4200), Capital, Santiago del Estero, Argentina
Tel.: (011) (0385) 424-0587
Fecha de trabajo recibido: 10/09/15
Fecha de trabajo aceptado: 10/11/15
Conflictos de interés: Franklin Ábalos, *speaker* de los laboratorios Novo Nordisk y Sanofi Aventis. El resto de los coautores declara que no existen conflictos de interés.

INTRODUCCIÓN

Cuenta una leyenda urbana que un supermercado en Estados Unidos se descubrió que, en sus ventas de los viernes, había una asociación entre: ¡compradores hombres, pañales y cerveza! ¿Cómo era esto posible? Aparentemente estos jóvenes padres llevaban pañales a su hogar y al mismo tiempo disfrutaban del fin de semana bebiendo cerveza[2]. Este hallazgo fue posible gracias a un proceso de minería de datos (MD), que en grandes bases de datos encuentra relaciones desconocidas o inesperadas.

Otro ejemplo: cuando adquirimos un libro en Amazon, el mismo sitio web nos sugiere: "Quienes compraron este mismo libro, mostraron también interés en los siguientes artículos…", y enumera los mismos con precios y características, etc. Estas grandes tiendas *online* lo hacen a través de potentes programas de MD, que a través de técnicas especiales (*clustering, association rules,* etc.), analizan al instante compradores históricos similares y sus preferencias.

La MD es un proceso que se encarga de descubrir conocimientos interesantes, como patrones y conexiones significativas en cualquier medio de almacenamiento de información.

La MD no es en sí un método estadístico. La estadística utiliza muestras y a partir de ahí hace

inferencias para toda la población. En cambio, la MD es una técnica que analiza voluminosos datos de variables y descubre tendencias, revela relaciones ocultas o asociaciones de los distintos atributos utilizando patrones automatizados. A través de ello, se puede clasificar, agrupar y predecir comportamientos en grupos o bases similares, y emplear dichos procesos a futuro. Esta búsqueda se lleva a cabo utilizando métodos de inteligencia artificial, matemáticos, estadísticos o algorítmicos.

El desarrollo de la MD tuvo estrecha relación, en sus comienzos, con el comercio y las finanzas. Retención, pérdidas o segmentación de clientes, fraudes bancarios, riesgo de préstamos, predicción de ventas, entre otras, fueron sus aplicaciones en los inicios[3].

Su utilización en el campo de la salud fue a posteriori, pero aún hoy está poco difundida. La MD es una novedosa técnica que viene desarrollándose en los últimos años, y es interesante conocer sus características, familiarizarse con sus términos y descubrir sus ventajas. Conocer *data mining* nos dará un plus en nuestros conocimientos científicos.

MÉTODOS

Existen básicamente dos tipos de técnicas en MD:

1. Supervisadas. Hay una variable que queremos pronosticar, y un set de variables que usamos para construir ese pronóstico (ejemplos: árboles de regresión, red neuronal artificial). Se necesita señalar en la base de datos cuál es la variable dependiente. Por lo tanto, las técnicas supervisadas permiten predecir el valor de una variable objetivo o dependiente, y aplicar dicho conocimiento en otros individuos.

Si la variable dependiente es numérica, tendremos que hacer un proceso de regresión o predicción; si la variable dependiente es categórica, corresponde una técnica de clasificación.

2. No supervisadas. No tienen variable dependiente (ejemplos: reglas de asociación, agrupamientos o *clustering*). Su tarea es descriptiva: buscan patrones que describen, asocian o agrupan datos. No pretenden predecir, no hay variable objetivo o dependiente.

Si nuestras variables son categóricas, haremos un proceso de asociación; si en cambio son numéricas, lo indicado es un proceso de agrupamiento.

Pasos en la minería de datos

• Hipótesis. ¿Cuál es el problema? ¿Cómo puedo investigar y resolverlo?

• Generación de la base de datos. No hay MD sin base de datos. En nuestro grupo de trabajo utilizamos hace muchos años el programa SPSS, con una base de datos *ad hoc* para pacientes con diabetes. Luego trasladamos esta información a un programa específico de MD (Rapidminer).

• Preparación de la base de datos. Es una de las fases más importantes, denominada "pre-procesado"; se traduce en garantizar la mayor fidelidad y prolijidad de los datos que se emplearán como materia prima para los análisis. Es la etapa más engorrosa y abarca aproximadamente el 70% del trabajo. En este paso se completan la identificación de variables, la edición de atributos, la modificación de etiquetas, se completa la depuración previa, el tratamiento de datos ausentes o erróneos, la importación de la matriz de datos, etc.

• Modelado. Es la aplicación de algoritmos (el proceso en sí mismo) para investigar, buscar y encontrar patrones relacionados de las variables en estudio. Se requiere el conocimiento de las distintas técnicas de MD y utilizar la correcta de acuerdo a nuestro objetivo.

• Evaluación. Determinar si se han encontrado patrones de comportamiento interesantes en nuestra base de datos que puedan responder a nuestras preguntas originales.

• Utilización. Esto es la aplicación del modelo en nuevas unidades de estudio.

Este trabajo muestra las características básicas de la MD, describiendo sus métodos y presentando algunas posibles aplicaciones en el campo de la diabetes. Las técnicas de MD son numerosas, algunas más sencillas y conocidas que otras. Como concepto general, con el manejo del 20% de ellas[4], las menos complicadas, se puede realizar el 80% de todos los procedimientos comunes.

Presentaremos tres ejemplos que no intentan ser trabajos científicos sino aplicaciones hipotéticas, con distintas técnicas, en temas relacionados con la diabetes:

1) Reglas de asociación en obesidad, sedentarismo y dislipidemia en pacientes con diabetes.

2) Agrupamiento *(clustering)* de pacientes con diabetes con distinto perfil lipídico.

3) Modelo matemático predictivo de eventos hipoglucémicos graves en pacientes con diabetes tipo 1, utilizando red neuronal artificial.

RESULTADOS

Reglas de asociación en relación a obesidad, sedentarismo y lípidos en pacientes con diabetes

Intentaremos ver cuáles son las asociaciones que ocurren con mayor frecuencia en una población de pacientes con diabetes en relación a lípidos, obesidad y sedentarismo.

Las reglas de asociación[5] se utilizan para descubrir relaciones o conexiones, que ocurren en común, dentro de un determinado conjunto de datos. Se usan con frecuencia en el comercio para detectar qué productos son adquiridos en forma conjunta (recordar los ejemplos de pañales y cerveza). Se denomina en inglés "*market basket analysis*", que podría traducirse en forma sencilla como "análisis del canasto del supermercado".

Existen tres factores que describen la frecuencia de asociaciones de dos o más variables: la confianza, el soporte y el *lift*.

- Confianza. Si observamos en 10 compras del supermercado que en cuatro se compró cerveza (premisa o antecedente) y en siete se compraron papas fritas (conclusión o consecuencia), y que en tres de las que se compró cerveza también se compraron papas fritas (3 de 4), la confianza para la regla cerveza-papas fritas es 75%. Se calcula simplemente al dividir la cantidad de veces que se compraron ambos artículos juntos por la cantidad de veces que se podrían haber comprado juntos.

Otra regla distinta sería papas fritas-cerveza. En este caso, aplicando el mismo criterio, la confianza es 43% (3 de 7). Significa que si un cliente ingresa en el comercio a comprar cerveza, es 75% probable que también compre papas fritas, pero si ingresa a comprar papas fritas, comprará cerveza en el 43% de las veces. La confianza mide la fortaleza de la regla de asociación.

- Soporte. Mide la frecuencia (del atributo o de la regla). Es el número de veces que un atributo o una regla ocurren, dividido por la cantidad de observaciones (frecuencia relativa). El soporte del atributo cerveza es 40% (4/10). El soporte de la regla cerveza-papa fritas es 30% (3/10).
- Lift. Ésta es una métrica interesante. Si el valor es 1 ó cercano a 1, significa que la relación es producto del azar. Si el lift está alejado de 1, indica una relación más fuerte, con más frecuencia de lo que indica el azar. Se obtiene así: soporte de la regla/soporte ítem A multiplicado soporte ítem B. En el caso de la regla cerveza-papas fritas: 0.75 / 0.7* 0.4 = 2.67.

Cuando fijamos los parámetros de nuestro modelo de reglas de asociación, podemos predeterminar el valor de confianza mínimo que deseamos, y el programa mostrará solamente las reglas que cumplan ese criterio. También podemos cambiar el nivel de soporte requerido.

Veamos los resultados de nuestro estudio de reglas de asociación.

Observamos una base de datos con 5.611 pacientes con diabetes tipo 2, en los que analizamos sus valores de colesterol total (CT) (mg%), LDL (low density cholesterol) colesterol (mg%), HDL (high density cholesterol) colesterol (mg%) y triglicéridos (TG) (mg%). También utilizamos las variables sedentarismo y obesidad. La técnica de reglas de asociación utiliza variables binominales, así que trasformamos nuestras variables numéricas en dicotómicas (colesterol mayor o menor de 200 mg%, LDL colesterol mayor o menor de 100 mg%, TG mayor o menor de 150 mg%, HDL colesterol mayor o menor de 40 mg% hombres/ 50 mg% mujeres, ausencia o presencia de sedentarismo, ausencia o presencia de obesidad). Exportamos la base preparada a Rapidminer, realizamos el pre-procesamiento correspondiente y el procesamiento con los operadores adecuados (Figura 1).

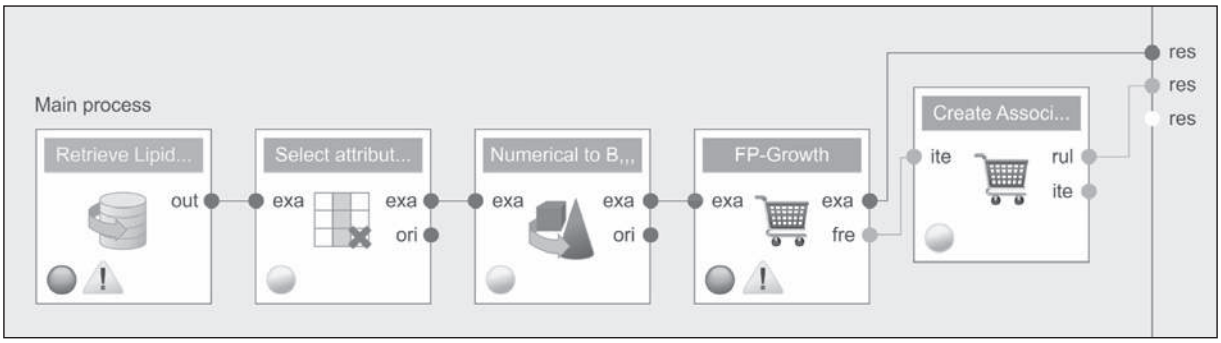

Figura 1: Aspecto de la pantalla de diseño de Rapidminer, con las etapas de importación de base de datos, preprocesamiento y procesamiento.
Fuente: elaboración propia.

• Soporte (algunos ejemplos): el 66,8% de nuestros pacientes es sedentario, 47,2% obeso, 54,3% tiene TG mayor a 150 mg%, 31,9% tiene TG mayor a 150 mg% combinado con LDL colesterol mayor a 200 mg%, 33,9% es sedentario y además obeso.

• Reglas de asociación. Se indican valores mínimos de soporte (20%) y confianza (70%). Se presentan al correr los resultados siete reglas de asociación. Ejemplo de regla de asociación: CT superior a 200 mg% combinado con TG superior a 150 mg% y LDL colesterol superior a 100 mg%: soporte 22,9% (frecuencia), confianza 70,8% (fortaleza). El lift es 2,219 lo que nos orienta a una "relación no azarosa". (Figura 2).

ASSOCIATION RULES (Create association rules)

No.	Premises	Conclusion	Support	Confid...	LaPl...	Gain	p-s	Lift	Convi...
1	Colesterol más de 200 mg	TG más de 150, LDL mayor a 100 mg	0.229	0.708	0.929	-0.417	0.126	2.219	2.333
2	TG más de 150, LDL mayor a 100 mg	Colesterol más de 200 mg	0.229	0.717	0.932	-0.410	0.126	2.219	2.391
3	Obesidad	Sedentarismo	0.339	0.718	0.909	-0.606	0.024	1.074	1.176
4	LDL mayor a 100 mg, Colesterol más de 200 mg	TG más de 150	0.229	0.741	0.939	-0.389	0.061	1.366	1.768
5	Colesterol más de 200 mg	TG más de 150	0.242	0.748	0.938	-0.405	0.066	1.378	1.814
6	TG más de 150, Colesterol más de 200 mg	LDL mayor a 100 mg	0.229	0.947	0.990	-0.255	0.108	1.898	9.438
7	Colesterol más de 200 mg	LDL mayor a 100 mg	0.309	0.955	0.989	-0.338	0.147	1.915	11.217

Figura 2: Aspecto de la pantalla con resultados de las reglas de asociación. Fuente: elaboración propia.

Se pueden modificar valores mínimos de soporte o confianza. Por ejemplo, si disminuimos soporte a 0,20 se mostrarán atributos y reglas con soporte desde el 20%.

Agrupamiento (*clustering*) de pacientes con diabetes de acuerdo a IMC (índice de masa corporal) y valores de colesterol total, LDL colesterol y triglicéridos

Clustering es una técnica muy útil para explorar datos, en especial cuando hay muchos atributos y no existe una separación natural de los casos. Este procedimiento pertenece al grupo de técnicas no supervisadas porque no hay una variable objetivo. Utiliza atributos numéricos. Un buen método de agrupamiento es aquel en que cada caso de un determinado clúster presenta características comunes con los otros integrantes del mismo y diferentes de los pertenecientes a los otros clústeres.

Nuestro grupo de trabajo en Santiago del Estero (Argentina) utiliza en forma rutinaria la determinación del riesgo cardiovascular de acuerdo al software de UKPDS (United Kingdom Prospective Diabetes Study), que calcula el riesgo a 10 años de desarrollar eventos cerebrales y cardíacos. Se hace en forma individual, uno por uno cada paciente. En nuestra base de datos podemos separarlos en grupos de acuerdo a esta medición de riesgo vascular.

Con la minería de datos podemos realizar esta separación en grupos de distintos riesgos, a través de la técnica de agrupamiento, incorporando las variables que deseemos (factores de riesgo). Se agrupan los pacientes en la cantidad de grupos que indiquemos (dos o más), de acuerdo a los valores de dichas variables.

Analizamos como ejemplo una base de datos con 5.611 pacientes con diabetes tipo 2, en los que observamos sus valores de colesterol total (mg%), LDL colesterol (mg%), triglicéridos (mg%) e índice de masa corporal. Realizamos el proceso para obtener tres grupos. Se realizó en el pre-procesamiento la normalización de las variables (porque presentaban unidades de medición distintas entre IMC y marcadores bioquímicos), y posterior al proceso de agrupamiento se "desnormalizó" para observar los datos con las unidades de origen en sus grupos definidos (Figura 3).

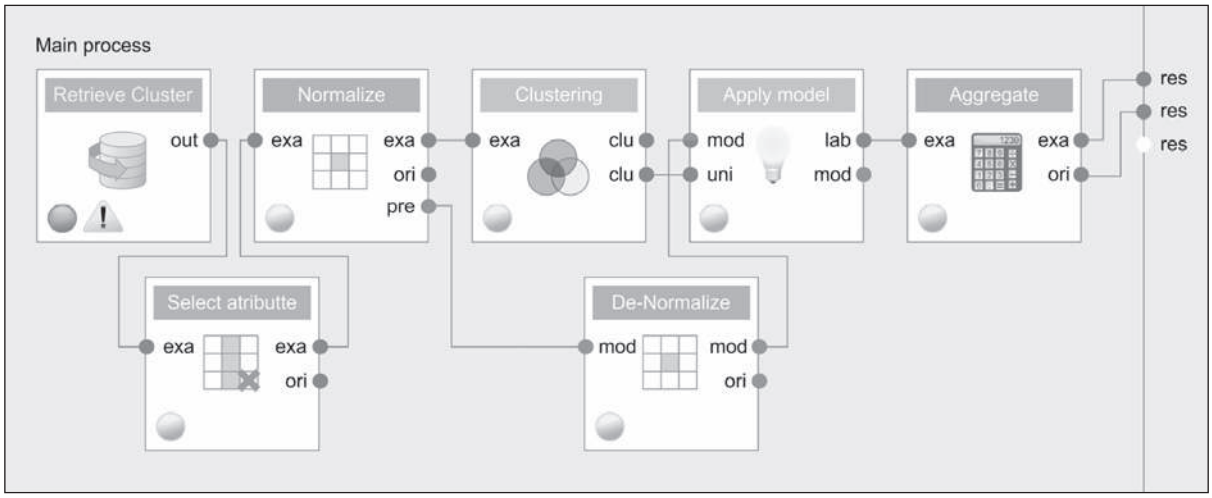

Figura 3: Pantalla de diseño del estudio de agrupamiento o clustering.
Fuente: elaboración propia.

El tratamiento de datos ausentes[6] puede variar de acuerdo a cada caso, se podrá ignorarlos, o utilizar un promedio, o el modo, u otro valor estadístico.

Se puede desplegar cada clúster y observar a cada ítem o paciente con su número de identificación y sus respectivas características (Figura 4).

EXAMPLE SET (5604 examples, 2 special attributes, 4 regular attributes)

Row No.	id	cluster	Indice de Masa Corporal	Triglicéridos...	Colesterol ...	LDL colesterol ...
1	1	cluster_1	30.900	251	234	38
2	2	cluster_2	27.300	98	408	291
3	3	cluster_1	25.700	286	221	45
4	4	cluster_1	33.300	160	173	101
5	5	cluster_1	36.500	138	215	56
6	6	cluster_1	25.800	134	206	43
7	7	cluster_2	32.200	429	222	56
8	8	cluster_2	18.100	75	303	188
9	9	cluster_2	28.900	550	211	51
10	10	cluster_1	36.600	163	183	52
11	11	cluster_1	41	156	177	95
12	12	cluster_2	31	483	208	62
13	13	cluster_1	35.300	535	193	28
14	14	cluster_0	29.400	197	215	100
15	15	cluster_1	39.200	155	176	69
16	16	cluster_0	30.200	127	239	128
17	17	cluster_1	33.900	86	177	80
18	18	cluster_1	38	84	164	65
19	19	cluster_2	31	342	255	169
20	20	cluster_0	21	76	281	155
21	21	cluster_0	22.200	78	217	122

Figura 4: Agrupamiento o clustering. Pacientes con número de identificación y valores normalizados de las variables estudiadas.
Fuente: elaboración propia.

Modelo matemático predictivo de eventos hipoglucémicos graves en pacientes con diabetes tipo 1 utilizando red neuronal artificial

La hipoglucemia severa (episodio hipoglucémico en el cual el paciente requiere asistencia de un tercero) es un evento agudo, potencialmente peligroso y difícil de predecir[7].

Realizamos un proyecto de estudio con MD investigando diferentes variables en pacientes con diabetes tipo 1, que tuvieron o no hipoglucemias severas (n:54 pacientes) para desarrollar un modelo predictivo aplicable en otros pacientes.

- *Variables en estudio*: edad, antigüedad de diabetes tipo 1, hemoglobina glucosilada A1c, frecuencia de automonitoreo de glucosa, porcentaje de insulina basal, unidades de insulina por kilogramo de peso, nivel de educación diabetológica, examen con monofilamento en pies, sedentarismo, conocimiento de recuento de hidratos de carbono, tabaquismo, alcoholismo.
- *Atributo objetivo*: antecedente de hipoglucemia severa.

Desarrollamos los pasos de MD:

1. Hipótesis de trabajo. Lo descripto.
2. Base de datos (importada desde SPSS). En este caso es conveniente "transformarla" en archivo Excel que nos permite manejar mejor la importación.
3. Pre-procesamiento. Cuando realizamos un estudio en forma prospectiva y diseñamos la base de datos específicamente para ello; esta etapa es menos engorrosa. Transformación de atributos nominales a numéricos.
4. Procesamiento. En este caso el operador se denomina "*neural net*".
5. Modelado. Observamos resultados (precisión).
6. Aplicación del modelo. Con un operador (*filter*) separamos pacientes que ya conocemos el resultado del desenlace (tuvieron o no hipoglucemia severa), de los que queremos evaluar. Podremos observar la predicción en nuevos pacientes y sus porcentajes de probabilidades de pertenecer a un grupo de menor o mayor chances de desarrollar eventos hipoglucémicos graves.

Accuracy: 51,33% +/- 21.56% (mikro: 51.85%) PerformanceVector (Performance)

	true Nunca	true Menos de 1 episodio/año	true Más de un epsodio/año	class precision
pred. Nunca	18	8	4	60.00%
pred. Menos de 1 episodio/año	11	4	1	25.00%
pred. Más de un epsodio/año	2	0	6	75.00%
class recall	58.06%	33.33%	54.55%	

Figura 5: Pantalla de resultados del estudio de predicción de hipoglucemia con la técnica de red neuronal artificial. Fuente: elaboración propia.

Observamos en esta tabla de contingencia que nuestro modelo predictivo tiene una baja precisión: 51,33%. Significa que en forma aproximada, sólo en la mitad de casos nuevos podríamos predecir a qué grupo pertenecerían (bajo, mediano o alto riesgo de sufrir hipoglucemia severa). Debemos mejorar nuestra precisión.

Habría diferentes caminos para mejorar nuestros resultados en el proceso:

- Agregar más pacientes a nuestra base de datos de referencia.
- Usar un modelo diferente (árbol de decisión, etc.).
- Cambiar datos o variables en el pre-procesamiento.
- Cambiar parámetros del modelo (número de "capas neuronales").
- Probar, ir para adelante y para atrás, experimentar, "picar piedras".
- Utilizar varios modelos a la vez (método ensamblado) con el operador "*vote*".

DISCUSIÓN

Algunas aplicaciones prácticas de los casos descriptos

Ejemplo 1 (diabetes, obesidad, sedentarismo y lípidos -reglas de asociación-)

• Regla colesterol total/LDL colesterol y triglicéridos elevados: porcentaje de pacientes con indicación terapéutica estatina-fibratos y evaluación de la controversia terapéutica de dicha asociación.

• Prevalencia del síndrome metabólico con distintas variables asociadas.

• Podríamos investigar: ¿qué dislipidemia se asocia con más frecuencia con sedentarismo en nuestros pacientes con diabetes? ¿Con qué frecuencia se presenta obesidad en nuestros pacientes hipertensos? ¿Qué porcentaje de pacientes con obesidad presenta hipertrigliceridemia? Las que se nos ocurran. Simplemente debemos "picar las piedras" y obtener las respuestas.

Ejemplo 2 (diabetes, índice de masa corporal y lípidos -agrupamiento-)

• En clúster de alto riesgo implementar medidas diferenciadas (educación, planes de actividad física, asesoramiento nutricional, acceso a medicación, etc.).

Ejemplo 3 (diabetes tipo 1 e hipoglucemia -red neuronal artificial-)

• Determinar el grupo de alto riesgo para hipoglucemia grave y reforzar las medidas preventivas (educación diabetológica, automonitoreo, provisión de glucagón, consejos a familiares, etc.).

Desde el punto de vista metodológico, aunque la MD nos ayuda a descubrir patrones y tendencias en los datos, no puede corregir errores estructurales en nuestra base de datos, ni explicar eventos y desenlaces, ni prometernos resultados perfectos.

Como todas las herramientas tecnológicas, la MD no puede desconocer la capital importancia del conocimiento clínico, la experiencia práctica y el análisis subjetivo frente a cada paciente en la toma de decisiones diagnósticas o terapéuticas.

Uno puede sentirse muy cómodo o feliz con una predicción de MD, pero si ésta no coincide con nuestro análisis clínico, debemos ser cautos en nuestro manejo a posteriori. La minería de datos no es un sustituto de la experiencia sino un complemento. Debemos trasladar nuestro modelo predictivo a la "vida real", revisarlo, y aceptar modificaciones en base al concepto clínico.

Data mining es un proceso, no una conclusión. No goza de infalibilidad y no reemplaza al ojo del médico. Es un desarrollo que ayuda a lograr conocimiento, entendimiento y aplicación del análisis de bases de datos en una forma singular.

Cómo iniciarse en minería de datos

Los programas de MD son variados. Los más utilizados son Rapidminer y R. El primero -Rapidminer- es más sencillo, gratuito y utiliza una interface gráfica que es amigable. El programa R es más complejo, se necesita aprender un lenguaje especial. Para los que no venimos de las ciencias informáticas, el empleo de Rapidminer es más recomendable.

¿Dónde aprender Rapidminer? Algunas recomendaciones para iniciados son:

• La página oficial de Rapidminer. Tiene un tutorial inicial excelente.

• Un excelente curso en www.bioestadístico.com, con videos (en castellano), del Dr. José Supo (Perú).

• Libros:

- Data mining for the masses.
- El manual oficial de Rapidminer.
- Exploring data with Rapidminer.
- Predictive analytics and datamining: concepts and practice with Rapidminer.

CONCLUSIONES

La minería de datos es una herramienta novedosa que permite el análisis de grandes volúmenes de datos, y descubre relaciones, conexiones y asociaciones entre ellos. A pesar de su escaso desarrollo, puede aplicarse en medicina.

Nuestra intención fue describir sus características básicas, utilizar algunos ejemplos prácticos relacionados con la diabetes para demostrar su eventual aplicación y estimular su conocimiento en colegas deseosos de complementar sus conocimientos en diabetología con técnicas nuevas de estudios de datos.

REFERENCIAS

1. Savage N. Mining data for better medicine. MIT Technology Review. 2011; 38:235-7.
2. North MA. Data mining for the masses. A global text project book. 2012; 122-24.
3. Kotu V, Deshpande B. Predictive analytics and data mining: concepts and practice with Rapidminer. Morgan Kaufmann, Elsevier, 2014: 4111.
4. Chisholm A. Exploring data with Rapidminer. Packt Publishing. 2013: 306-7.
5. Hofmann M. Data mining use cases and business and analytics applications. Chapman&Hall CRC. 2013; 234-5.
6. Sumathi S, Sivanandam S. Introduction to data mining and its applications. Springer. USA. 2006; 732-33.
7. The DCCT Research Group. Epidemiology of severe hypoglycemia in the diabetes control and complications trial. The American Journal of Medicine. Vol. 90, Issue 4, April 1991, Pages 450-459.