

Identificación probabilística de interacciones medicamentosas

Luis Enrique Colmenares Guillén¹, Luis Daniel Oidor Juárez², José Gustavo López y López³

^{1, 2}Facultad de Ciencias de la Computación

³Facultad de Ciencias Químicas

Benemérita Universidad Autónoma de Puebla

Puebla, México

¹lecolme@gmail.com

²daniel_oidor@hotmail.com

³jose.lopez@correo.buap.mx

Resumen. La fabricación de productos farmacéuticos representa económicamente, una de las industrias más importantes en todo el mundo. Actualmente, existen tratamientos para casi cualquier enfermedad conocida y con más de un fármaco para cada una de estas. Esto da como resultado que un paciente deba administrarse varios medicamentos durante la terapia farmacológica, con el riesgo consecuente de que existan interacciones negativas para el organismo entre los mismos, denominadas interacciones medicamentosas. El presente trabajo pretende diseñar un método basado en un algoritmo clasificador para identificar las interacciones medicamentosas, mediante el procesamiento de grandes cantidades de datos de forma automatizada. El algoritmo diseñado ayudará a la implementación de sistemas de apoyo para la actividad relacionada con la idoneidad de la prescripción médica. El método que se propone tiene posibilidades de incrementar su eficiencia de forma gradual, a través de conjuntos de entrenamiento, lo que lo convierte en una solución rápida, flexible y adaptable.

Palabras Clave: algoritmo; automatización; clasificador; cómputo; corpus; idoneidad; interacción; farmacología; medicamentos; prescripción.

1. Introducción

Desde tiempos ancestrales, el ser humano se encuentra en una búsqueda permanente de la curación de sus enfermedades, desde el inicio de la civilización, en que sus males eran atribuidos a seres malignos y hechos mágicos. En nuestros días, utilizamos las herramientas que la ciencia y la tecnología para la cura de enfermedades.

Hasta la primera mitad del siglo pasado, el hombre utilizó remedios que, en su gran mayoría, no alteraban de forma importante los mecanismos fisiológicos. Posteriormente la medicina cambió, introduciendo una inmensa gama de medicamentos, capaces de modificar de manera favorable el curso de las enfermedades y la aparición de síntomas y signos. El papel que juegan los

medicamentos en las sociedades actuales es tan relevante que hoy en día, la industria farmacéutica es una de las más dinámicas e importantes para la economía mundial.

Ante este panorama, cada medicamento desarrollado debe seguir un minucioso proceso de pruebas para el aseguramiento de la eficacia del mismo, no sólo en términos de su calidad farmacéutica, sino también en función de la gravedad de los efectos secundarios y de las reacciones que estos puedan provocar en el ser humano.

Sin embargo, en la prescripción médica la indicación de uso de un solo medicamento es inusual, por lo general un paciente debe administrarse dos o más. Es por ello que los desarrolladores de medicamentos deben realizar también pruebas de combinaciones de estos, de modo que sea posible evitar una interacción medicamentosa que afecte de manera negativa al paciente, ya sea por inhibición de los efectos de uno de ellos, la generación de efectos adversos o el aumento de toxicidad de alguna de las sustancias activas.

Para ello, se han realizado múltiples estudios por parte de laboratorios, centros de investigación y empresas privadas en las áreas de química y medicina.

Pero, como se ha dicho antes, esta industria es una de las más grandes de la economía mundial. ¿Se podría aseverar que el médico conoce (y recuerda) todas y cada una de las posibles interacciones entre medicamentos que pudieran afectar nuestra salud? Definitivamente no. A pesar de los grandes avances tecnológicos, la prescripción médica es una actividad que sigue siendo aplicada por el profesional de la salud de forma “manual”, a través del uso de los conocimientos generados por la investigación científica en materia farmacéutica, sin asistencia de dispositivos o mecanismos automáticos, como ya sucede en otras áreas de la medicina como los análisis clínicos y el diagnóstico médico. Además actualmente se han desarrollado procesos que implican evaluación del perfil farmacoterapéutico con el fin de identificar problemas relacionados con los medicamentos (Idoneidad de la prescripción y Conciliación de la medicación), que mediante un proceso de intervención antes de la aplicación se evitan errores de medicación en el paciente hospitalizado.

Es necesario, entonces, diseñar métodos basados en el cómputo automático de grandes cantidades de datos, que sirvan como base para el desarrollo de herramientas que provean a los profesionales de la salud y a los desarrolladores de fármacos, apoyo en la identificación de las interacciones medicamentosas. Estos métodos deberán ser diseñados para explotar la información generada por los estudios e investigaciones realizados por los especialistas en la materia, siendo su principal aporte la automatización del procesamiento de información y la aplicación de la ciencia computacional en el desarrollo de soluciones que potencien la actividad humana.

Identificar correctamente las interacciones medicamentosas es un proceso complejo, dado que el resultado de calificación de la interacción dado un par de medicamentos específico generalmente es un conjunto de posibles efectos adversos provocados por ese par de medicamentos. Cada uno de estos efectos tendrá una probabilidad de suceder en un paciente específico.

Dada esta complejidad, el algoritmo que se diseñará será un clasificador, con la finalidad de que nos proporcione un conjunto de resultados que puedan ser ponderados conforme a su probabilidad.

El objetivo principal de este trabajo es proponer una aproximación a un algoritmo clasificador que permita la identificación y ponderación probabilística de

interacciones medicamentosas. El diseño del algoritmo estará orientado a que la identificación se lleve a cabo de forma previa a la prescripción médica, o bien como una herramienta para el proceso de la Idoneidad de la prescripción, a fin de detectar problemas relacionados con los medicamentos y evitar resultados negativos de la medicación.

Para alcanzar este objetivo será necesario llevar a cabo algunas tareas que conforman los objetivos específicos del presente trabajo, las cuales se enumeran a continuación.

- Realizar una investigación de los métodos actuales para la identificación de interacciones medicamentosas.
- Identificar las fuentes confiables de información sobre las características, efectos y contraindicaciones de medicamentos.
- Analizar las ventajas y características generales de los diferentes tipos de clasificadores.
- Determinar el tipo de clasificador en que se basará el diseño del algoritmo que será producto del proyecto.

Para alcanzar este objetivo será necesario llevar a cabo algunas tareas que conforman los objetivos específicos del presente trabajo, las cuales se enumeran a continuación.

2. Fuentes de información sobre medicamentos

La información es un elemento poderoso y fundamental en el desarrollo de cualquier actividad humana de cualquier índole. En el caso de la medicina, y en específico de la prescripción médica, contar con la información correcta y oportuna es un requisito indispensable.

Para que la ciencia computacional pueda ofrecer herramientas que coadyuven con el objetivo de la medicina es necesario que ésta última le proporcione fuentes de información que sirvan como base de dichas herramientas.

En la actualidad, el principal problema es seleccionar la información más relevante y de mayor calidad [1]. Toda la información relacionada con los medicamentos está incluida en un campo más amplio, conocido como información biomédica.

2.1. Pirámide de Hynes

Las fuentes de información biomédicas (incluyendo lo referente a medicamentos) pueden clasificarse mediante el modelo piramidal propuesto por **R. Brian Haynes**, mejor conocido como *Pirámide de las 5s*, que puede observarse en la Fig. 1.

El modelo de Haynes es el más aceptado para la clasificación de fuentes de información relacionadas con la biomedicina. En él se pueden diferenciar cinco niveles que a continuación se describen brevemente.

- Estudios. Representan las fuentes primarias, incluyen los artículos biomédicos originales.
- Síntesis. Incluyen recursos que indexan y publican revisiones sistemáticas y metaanálisis.
- Sinopsis. Comprenden resúmenes estructurados de artículos originales, así como boletines e informes de evaluación de medicamentos elaborados por comunidades autónomas, hospitales y agencias reguladoras.
- Compendios. Incluyen revisiones sistemáticas y resúmenes sobre patologías o tratamientos determinados, constituyen las fuentes que eran clasificadas en la forma clásica como secundarias.
- Sistemas. Aquí se incluyen aplicaciones de cómputo auxiliares en la toma de decisiones, como bases de datos y sistemas de evaluación automatizados.

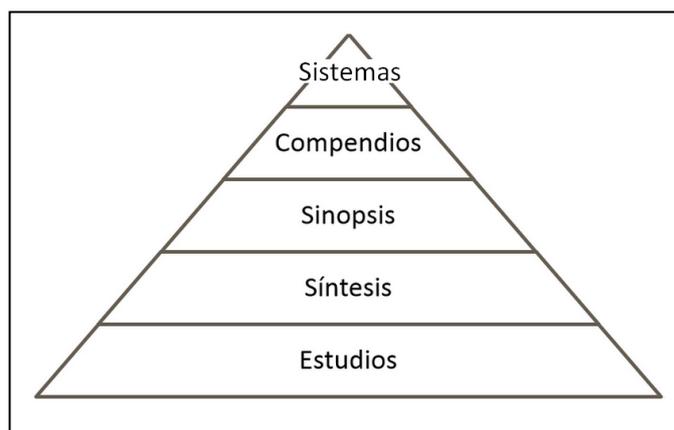


Fig. 1. Modelo piramidal de las 5s de Haynes.

2.1 Principales fuentes de información sobre medicamentos

Como parte del proceso de investigación comprendido en el presente trabajo, se identificaron tres principales fuentes de información referentes a medicamentos. Estas fuentes fueron elegidas por ser las más aceptadas en el campo de la investigación biomédica.

Diccionario de Especialidades Farmacéuticas de Thomson. Mejor conocido como PLM, es una de las fuentes más utilizadas por los médicos y farmacias en México. Incluye una lista completa de todos los fármacos aprobados por la Secretaría de Salud en México, lo que conforma un total de más de 3,500 productos

farmacéuticos. Se encuentra disponible en su versión impresa y, gracias a los avances tecnológicos actuales, en versiones digitales en Internet y en forma de aplicaciones para dispositivos móviles. PLM México tiene a disposición de cualquier persona con acceso a Internet una versión web disponible en la dirección electrónica <http://www.medicamentosplm.com/> [2].

Micromedex. Es una base de datos norteamericana que contiene amplia información de medicamentos y sustancias relacionadas, así como de pruebas de laboratorio e interacciones medicamentosas. Su contenido es actualizado constantemente mediante la revisión sistemática de estudios médicos, siendo así una fuente confiable de información. Consta de diversos productos, entre los cuales se encuentra DRUGDEX, un sistema de información específico sobre medicamentos, su administración, efectos adversos e interacciones [3].

Vademécum. Es un catálogo de especialidades, entre las cuales se encuentra la información relacionada a los medicamentos. Puede encontrarse en versión impresa o en versión digital, que es distribuida por medio de un CD-ROM que contiene todos los archivos necesarios para ser instalado en un equipo de cómputo. Contiene información relevante relacionada con productos farmacéuticos, y está destinado a los profesionales de la salud, tales como médicos y farmacéuticos [4].

Existen otras muchas fuentes de información, sin embargo se mencionan solamente las tres anteriores dada su alta aceptación y amplio uso por parte de los profesionales de la salud. La identificación de estas fuentes como fiables y generalmente aceptadas es un paso fundamental en la elaboración de un método para la identificación de interacciones medicamentosas. Un método basado en información de calidad estará en condiciones de proporcionarnos resultados de calidad, como los que son requeridos en la práctica de la prescripción médica.

3. Estado del arte

Durante las últimas décadas, la ciencia computacional se ha desarrollado rápida y continuamente, lo que ha producido computadoras de propósito general con mayor procesamiento de cómputo y movilidad. Así, la ciencia computacional se ha convertido en una herramienta de apoyo en la resolución de problemas.

Como parte de esta incorporación de las nuevas tecnologías a ámbitos de la actividad humana, la medicina y la farmacia han encontrado en la computación una herramienta poderosa para la agilización y mejoramiento de los procesos de apoyo biomédico.

Con relación a las interacciones medicamentosas, con ayuda de la ciencia computacional se han obtenido grandes avances, entre los que destacan algunos métodos diseñados con la finalidad de ayudar a la identificación de las causas de un efecto adverso, entre las cuales se pueden encontrar interacciones de medicamentos. A continuación, se mencionan algunos algoritmos más importantes.

Algoritmo de Karch y Lasagna. Es un algoritmo que, aunque ya han pasado más de 35 años de su publicación, sigue siendo un estándar para la identificación de los efectos adversos de medicamentos, entre los cuales existen combinaciones de dos o más fármacos. Contempla la secuencia temporal entre el cuadro clínico que presenta un paciente y los fármacos presuntos responsables de dicha sintomatología, mediante la evaluación de la relación causa-efecto. Esta relación puede clasificarse como Definida, Probable, Posible o Condicional [5].

Algoritmo de Kramer. Consiste en una secuencia de preguntas y una escala de calificación que permite, al final de la aplicación del cuestionario, establecer la causalidad por categorías. Consta de 56 preguntas dicotómicas (sí/no). Es también un algoritmo diseñado para determinar si una reacción fue generada por un medicamento o un conjunto de medicamentos en específico [6].

Algoritmo de Naranjo y colaboradores. Es un algoritmo basado en el de Karch y Lasagna y que consta de un cuestionario, como el algoritmo de Kramer, pero de menor cantidad de cuestiones (solamente 10 preguntas dicotómicas). Al igual que los algoritmos anteriores, no fue diseñado específicamente para determinar la interacción entre fármacos, sino de relacionar un efecto adverso con su causal. Resulta eficaz dada su simplicidad y su corta extensión [7].

La particularidad que comparten los tres métodos mencionados es que su objetivo es identificar de manera general la causa de un efecto adverso en un paciente, lo que nos hace pensar en ellos como métodos *a posteriori* con relación a la prescripción médica. Esto implica que una interacción medicamentosa deberá presentarse al menos una vez para poder ser evaluada e identificada como potencialmente negativa.

Un esfuerzo importante lo representa el **Corpus Drug-Drug Interactions**. Fue desarrollado a partir de una colección de textos médicos en los cuales figuran las diversas interacciones medicamentosas, incluso las que no han sido probadas por la ciencia médica, con su catalogación de certeza. Entre las fuentes de información biomédica utilizadas en el desarrollo de dicho corpus, se encuentra la base de datos **Micromedex**. Este trabajo genera las bases para realizar métodos de explotación de dicha información mediante la aplicación de técnicas de tratamiento de texto.

Un método de identificación de interacciones mediante la utilización de un corpus permitirá establecer la probabilidad de que esta se presente en un paciente, antes de que los medicamentos implicados sean prescritos, evitando así incomodidades y reacciones no deseadas en seres humanos.

4. Técnicas de clasificación

Para que la información contenida en un corpus sea procesada y clasificada es necesario diseñar un método basado en técnicas de clasificación de grandes cantidades de texto.

Se puede formalizar la *clasificación* como una aproximación de una función objetivo no conocida que describe la forma en la que instancias del problema deben

ser clasificadas, mediante otra función, denominada *clasificador*. La función objetivo se puede representar como en (1), mientras que el clasificador se encuentra representado en la forma mostrada en (2).

$$\Phi: I \times C \rightarrow \{T, F\} \quad (1)$$

$$\Theta: I \times C \rightarrow \{T, F\} \quad (2)$$

C es un conjunto predefinido de categorías, en tanto que I es un conjunto de instancias del problema. Es común representar cada instancia $i_j \in I$ como una lista $A = \{a_1, a_2, \dots, a_{|A|}\}$ de valores característicos, denominados atributos. Si $\Phi(i_j, c_i) = T$, entonces i_j es un ejemplo positivo de la categoría c_i . Si, por el contrario, $\Phi(i_j, c_i) = F$, entonces i_j es un ejemplo negativo de c_i .

Es posible generar de forma automática el clasificador mediante el proceso llamado *aprendizaje supervisado*. Este proceso implica la observación de los atributos de un conjunto de instancias ya clasificadas, de modo que sea posible asignar una instancia no clasificada en una determinada categoría. Un requisito para la construcción del clasificador es contar con una colección Ω , denominada *conjunto de entrenamiento*, de ejemplos tales que el valor de la función $\Phi(i_j, c_i)$ sea conocido para cada $(i_j, c_i) \in \Omega \times C$.

Existen diversas técnicas de clasificación, con diferentes fortalezas y características. A continuación, se revisan cuatro técnicas cuyos resultados en la clasificación de texto destacan sobre otras.

4.1. Vecinos más cercanos

Es un clasificador supervisado, también conocido como *k-NN*. Para su aplicación se exploran los atributos de los elementos del conjunto de entrenamiento para determinar la categoría a la que pertenecerá una nueva instancia no clasificada. Esta técnica utiliza la información suministrada por las k instancias del conjunto de entrenamiento más cercanas a la instancia que se desea clasificar.

En la Fig. 2 están representadas doce elementos pertenecientes a dos categorías diferentes, la *Categoría A*, conformada por seis cuadros de color verde, y la *Categoría B*, que la forman seis círculos de color naranja. La instancia x es la que deseamos clasificar. Dentro del círculo mayor podemos observar los tres vecinos más cercanos, es decir, que $k = 3$.

Dado que, de los tres vecinos más cercanos, uno pertenece a la *Categoría A* y dos a la *Categoría B*, después de aplicar la técnica, la instancia x será asignada a la *Categoría B*.

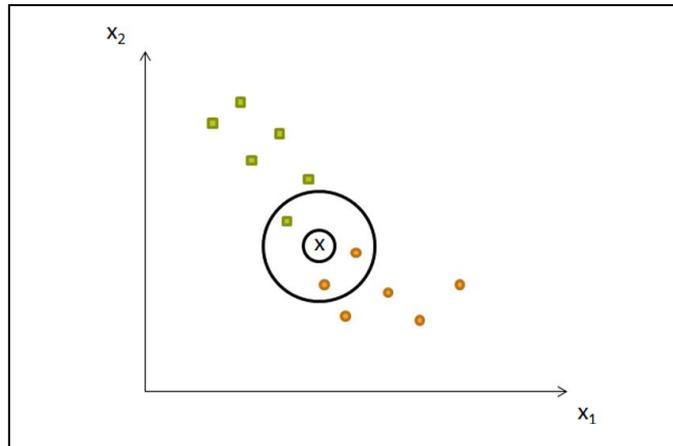


Fig. 2. Técnica de clasificación de los vecinos más cercanos.

4.2. Naïve Bayes

Es un clasificador probabilístico basado en el **Teorema de Bayes**, enunciado por **Thomas Bayes**. Básicamente, este teorema relaciona la probabilidad de un evento A dado B con la probabilidad del evento B dado A .

La fórmula del **Teorema de Bayes** es:

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^n P(A_i)P(B|A_i)}$$

Es importante recordar la siguiente definición de la probabilidad condicional:

$$P(BA_i) = P(B|A_i)P(A_i) \quad (3)$$

El método de **Naïve Bayes** usa frecuencias para calcular probabilidades condicionales con el fin de realizar predicciones sobre nuevas instancias del problema. Un clasificador de este tipo puede ser tanto descriptivo como predictivo. A continuación se formalizará la definición del clasificador Naïve Bayes.

Dados los eventos E y F , se sabe que:

$$E = EF \cup EF^c$$

EF y EF^c son mutuamente excluyentes, por lo tanto:

$$P(E) = P(EF) \cup P(EF^c)$$

$$P(E) = P(E|F)P(F) + P(E|F^c)P(F^c)$$

$$P(E) = P(E|F)P(F) + P(E|F^c)(1 - P(F)) \quad (4)$$

Lo que la ecuación (4) pone de manifiesto es que la probabilidad de E es una ponderación de su probabilidad condicional dada la ocurrencia y la no ocurrencia de F . Ahora, se supone que los eventos F_1, F_2, \dots, F_n son mutuamente excluyentes, lo que se puede expresar como:

$$E = \bigcup_{i=1}^n E_i$$

A partir de la definición de probabilidad condicional, expresada en (3) y dado que los eventos EF_i para todo i de 1 a n , son mutuamente excluyentes, se infiere que:

$$P(E) = \sum_{i=1}^n P(E F_i)$$

$$P(E) = \sum_{i=1}^n P(E|F_i)P(F_i) \quad (5)$$

De esta forma, la ecuación (5) muestra que, para los eventos F_1, F_2, \dots, F_n , de los cuales puede darse la ocurrencia de uno y solamente uno, se establece que la probabilidad de E es igual al promedio de las ponderaciones de $P(E|F_i)$.

La suposición de que los eventos son mutuamente excluyentes es la que da origen al nombre de **Naïve** (ingenuo), ya que esto no siempre sucede. A pesar de ello, el método ha sido implementado con buenos resultados, por lo que es uno de los clasificadores más aceptados.

4.3. Support vector machines

Las **Máquinas de Soporte Vectorial** o **Support Vector Machines (SVM)** son un conjunto de algoritmos cuya técnica se basa en el aprendizaje de dos categorías distintas de entrada. Dado que el objetivo es clasificar instancias en una sola categoría, con la descripción proporcionada por los datos elabora una frontera de decisión alrededor de los datos de aprendizaje, para luego buscar la separación máxima entre categorías. De esta forma se divide al espacio muestral en categorías distintas.

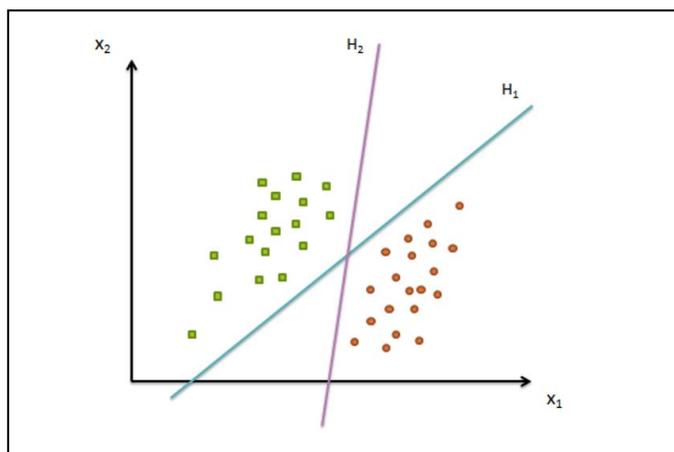


Fig. 3. Funcionamiento del algoritmo SVM.

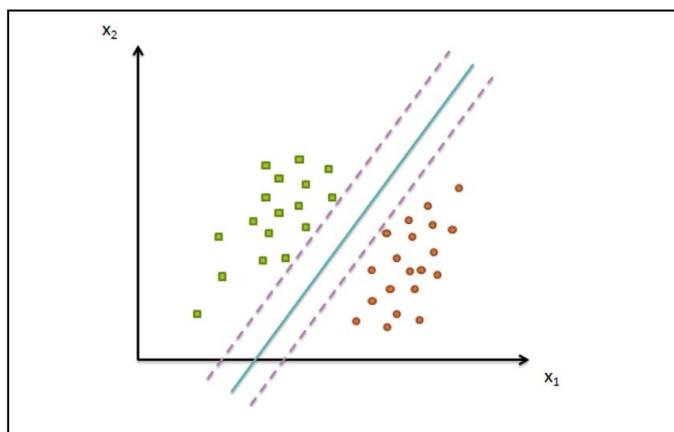


Fig. 4. Hiperplano y vectores de soporte para SVM.

En la Fig. 3, se muestra un ejemplo del funcionamiento de SVM. En él, se representan los datos en el plano xy . El algoritmo intentará encontrar un hiperplano de dimensión 1 que constituye el límite que separa las dos categorías existentes. La posición de una nueva instancia a un lado u otro de este hiperplano determinará la pertenencia de dicha instancia a la categoría específica correspondiente.

Como se puede observar, existe un número infinito de hiperplanos posibles que dividan las instancias en dos categorías, por lo que debe encontrarse el mejor de ellos. Para esto, el algoritmo elige el hiperplano cuyo margen de separación entre las instancias de ambas categorías sea el máximo, es decir, existirán dos líneas paralelas al hiperplano que indicarán la distancia con las instancias de cada categoría. A estas líneas se les llama *vectores de soporte*. En la Fig. 4 los vectores de soporte aparecen como líneas punteadas

5. Búsquedas mediante corpus

La definición más simple de **corpus** nos refiere a éste como una colección, generalmente amplia, de textos. Sin embargo, cuando el término es usado en el ámbito de la lingüística computacional, éste tiene más implicaciones.

La primera implicación se refiere al lugar donde este conjunto de textos está almacenado. Leech [8] introduce el concepto de corpus como un *emocionante fenómeno, una magnífica gran cantidad de texto, almacenada en una computadora*.

Por su parte, Francis [9] agrega a su definición de corpus el que esta colección *se asume como representativa de un determinado idioma, dialecto o subconjunto de un idioma para ser usado en análisis lingüístico*.

Sin embargo, tal vez la mejor definición la proporciona el grupo de trabajo que está dedicado a los corpus de texto. Denominado como EAGLES (Expert Advisory Group on Language Engineering Standards) [10], este grupo define un corpus como *una colección de piezas de un idioma seleccionadas y ordenadas de acuerdo a criterios lingüísticos explícitos con el fin de ser usados como ejemplo de un idioma*.

En el campo de la ciencia biomédica, se han desarrollado diversos corpus con el objeto de analizar lingüísticamente la información contenida relacionada con los aspectos biomédicos. Uno de los objetivos del presente trabajo es desarrollar un corpus que contenga información específicamente relacionada con las interacciones entre medicamentos.

La finalidad es generar un corpus de dominio biomédico orientado a la identificación de interacciones medicamentosas, que sea base para el desarrollo de un algoritmo clasificador que permita identificar interacciones entre medicamentos.

6. Aproximación a un algoritmo clasificador de interacciones medicamentosas

Hasta ahora, se ha proporcionado la base teórica sobre la cual se desarrollará una aproximación a un algoritmo clasificador de interacciones medicamentosas, basado en búsquedas a través de corpus.

Se sabe que los efectos de una interacción medicamentosa negativa pueden presentarse para ciertos pacientes, en tanto que en otros, éstos podrían ser imperceptibles o, incluso, nulos. El conocimiento o ignorancia sobre todas las causas que se relacionan con un efecto adverso hacen que la clasificación de éstos sea una función probabilística que pondere su ocurrencia para un paciente cualquiera.

Evidentemente, se intenta resolver un problema para el que se cuenta con datos parciales, solamente considerando la presencia de medicamentos, sin contemplar el resto de variables relacionadas con el paciente, como su peso, masa muscular, enfermedades crónicas o antecedentes familiares, entre muchos otros.

Por otro lado, es importante destacar que uno de los principales aportes de este trabajo es proporcionar las bases para el desarrollo de sistemas y herramientas de apoyo biomédico, en específico, para la prescripción médica. Actualmente, la identificación de una interacción medicamentosa por parte de un profesional de la salud, en el supuesto de que se desconozca o exista incertidumbre sobre la misma, comprende la búsqueda de la información de cada medicamento involucrado, proceso que se realiza sin ayuda de sistemas expertos.

Se busca, mediante el desarrollo de este método de clasificación, proporcionar una respuesta rápida de la probabilidad de interacciones entre medicamentos, lo que significará una ventaja sobre el proceso realizado de forma manual.

Dado lo anterior, se eligió el clasificador de **Naïve Bayes** como el tipo de algoritmo que se desarrollará, debido a la fortaleza de la técnica para realizar predicciones a partir de datos parciales y por su rapidez.

El desarrollo del presente trabajo generará beneficios a mediano y largo plazo, posibilitando el desarrollo de sistemas comerciales que puedan ser implementados en hospitales, farmacias y centros de salud, en los cuales se implemente el algoritmo desarrollado.

El algoritmo clasificador diseñado podrá operar bajo un esquema funcional como el que se muestra en la Fig. 5. Los datos de entrada, lo conforman un par de medicamentos, cuya posible interacción será evaluada por el algoritmo clasificador. Estos datos son ingresados mediante una terminal y llegarán a un servidor de aplicaciones, donde residirá la implementación del algoritmo. Desde ahí, se realiza la búsqueda y extracción de los datos relacionados en el corpus que contiene la información de cada medicamento para ser clasificados por el algoritmo. El resultado se envía a la terminal, que muestra si existe una posible interacción entre los medicamentos proporcionados.

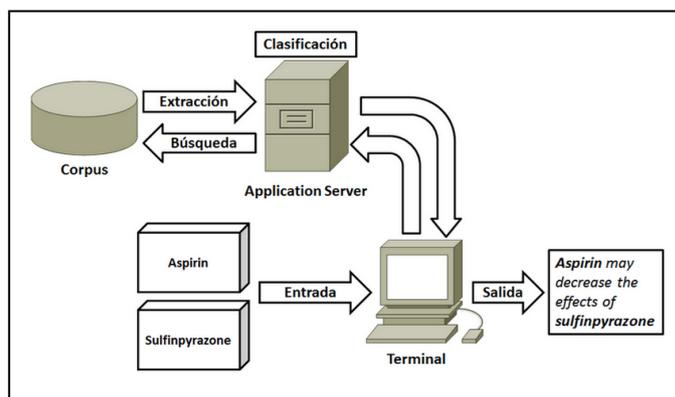


Fig. 5. Esquema funcional de una implementación del algoritmo clasificador.

Una característica importante de un clasificador es que su salida depende únicamente de la entrada de datos, en este caso, del corpus suministrado conteniendo la información sobre interacciones medicamentosas. Es decir, se trata de un método

incremental, que mediante la aplicación constante del algoritmo, a través de sistemas que lo implementen, generará más datos que pueden ser fácilmente incorporados al método original como entradas, con lo cual se expande el dominio de aplicación del algoritmo a mayores sectores poblacionales diferenciados por condiciones patológicas, de raza o con determinadas condiciones de salubridad.

7. Pruebas

Para comprobar la efectividad del método propuesto y, eventualmente, realizar adecuaciones y ajustes para obtener mejores resultados, se deberán realizar una serie de pruebas que nos ayuden al aseguramiento de la calidad del algoritmo clasificador.

Para tal efecto, se realizarán pruebas con cincuenta medicamentos de los más utilizados en México. El Laboratorio de Farmacia Clínica de la Benemérita Universidad Autónoma de Puebla ha facilitado una lista con medicamentos propuestos para ser considerados en las pruebas una vez terminado el desarrollo del algoritmo propuesto.

Una de las principales características de estas pruebas es que deberán realizarse en un ambiente real controlado, es decir, se buscará una unidad de servicios médicos para incorporar el uso del algoritmo en las prescripciones médicas que se realicen en dicha unidad. En esta etapa, los resultados que arroje el algoritmo conformarán un conjunto de casos que deberán ser evaluados por los profesionales de la salud encargados del control, administración y almacenamiento de medicamentos, para corroborar la efectividad del algoritmo.

El término controlado hace referencia al hecho de que la implementación que se desarrollará para efecto de las pruebas estará disponible solamente para llevar a cabo la evaluación de resultados, sin incorporarse por completo a la operación habitual de la unidad de servicios médicos donde se efectuarán las pruebas.

8. Conclusiones

La identificación de interacciones medicamentosas es uno de los problemas, dentro de la actividad de la prescripción médica, que requieren la aplicación de la ciencia computacional para desarrollar herramientas automatizadas que proporcionen apoyo a los profesionales de la salud.

Mediante el análisis del problema y el diseño de una propuesta de solución, se concluye que el presente trabajo es factible, dadas las condiciones actuales de las actividades relacionadas con la prescripción de medicamentos.

Como ya se explicó, la solución propuesta es un algoritmo clasificador basado en el de **Naïve Bayes** cuyos resultados serán ponderaciones de probabilidad de la posible interacción entre dos medicamentos.

El desarrollo del algoritmo permitirá su implementación en herramientas y sistemas para ser integrados a soluciones que permitan a los profesionales de la salud contar con información suficiente de apoyo en la toma de decisiones. Sin embargo, el principal impacto del presente trabajo es desarrollar un método para identificar interacciones medicamentosas previas a la prescripción médica, con la finalidad de evitar en el paciente efectos no deseados y afectaciones a su salud.

Referencias

1. Rancaño, I.; Rodrigo, J. A.; Villa, R.; Abdelsater, M.; Díaz, R.; Álvarez, D.: Evaluación de las páginas web en lengua española útiles para el médico de atención primaria, *Aten Primaria*, vol. 31, no. 6, pp. 575-584 (2003)
2. Diccionario de Especialidades Farmacéuticas 2012. <http://www.medicamentosplm.com>. Accedido el 25 de junio de 2013
3. Micromedex. <http://www.micromedex.com/>. Accedido el 26 de junio de 2013
4. PR Vademécum México. <http://mx.prvademecum.com/>. Accedido el 26 de junio de 2013
5. Armijo, J.; González, M.: Estudios de seguridad de medicamentos: Métodos para detectar las reacciones adversas y la valoración de la relación causa-efecto, *El ensayo clínico en España*, pp. 161-190 (2001)
6. Kramer, M. S.; Levental, J. M.; Hutchinson, T. A.; Feinstein, A. R.: An algorithm for the operational assessment of adverse drug reactions, I: background, description, and instructions for use, vol. 242, pp. 623-632 (1979)
7. Naranjo, C. A.; Busto, U.; Sellers, E. M.: A method for estimating the probability of adverse drug reactions, *Clin Pharmacol Ther*, vol. 30, pp. 239-245 (1981)
8. Leech, G.: Corpora theories of linguistic performance. J. Svartvik (Ed.): *Directions in Corpus Linguistics*, pp. 105-122 (1992)
9. Francis, W. N.: Problems Assembling and Computerizing Large Corpora. Johansson, S (Ed.), pp. 124-136 (1982)
10. Expert Advisory Group on Language Engineering, Text Corpora Working Group Reading Guide. EAG-TCWG-FR-2. (1996)
11. Freer, E. B.; Chavarria, J. C.: El desarrollo de la computación y su influencia en la medicina, *Revista Costarricense de Ciencias Médicas*, vol. 13, pp. 59-70 (1992)
12. Martín, H.; Martín, S.: Cómo localizar la mejor evidencia científica, *Recursos de información para la MBE* (2007)
13. Calderón, C. A.; Urbina, A. P.: La Farmacovigilancia en los últimos 10 años: actualización de conceptos y clasificaciones. Logros y retos para el futuro en Colombia, *Revista de los estudiantes de medicina de la Universidad Industrial de Santander*, vol. 24, pp. 57-73 (2011)
14. Mitchell, T.: *Machine learning*. Ed. Mc Graw-Hill (1997)
15. Mladenić, D.; Grobelnik, M.: Feature selection for unbalanced class distribution and Naïve Bayes. Department of Intelligent Systems, J. Stefan Institute (1999)
16. Pacheco, S. D.; Díaz, L. G.: El clasificador Naïve Bayes en la extracción de conocimiento de bases de datos, *Ingenierías*, Abril-Junio, vol. 27, no. 8 (2005)
17. Cortés, C.; Vapnik, V.: Support-Vector Networks, *Machine Learning*, vol. 20, pp. 273-297 (1995)
18. Aas, K.; Eikvil, L.: Text categorization: A survey, Norwegian Computing Center, *Technical Report* (1999)
19. Segura, I.: Application of Information Extraction techniques to pharmacological domain: Extracting drug-drug interactions, Tesis doctoral, Universidad Carlos III, España (2010)
20. Téllez, A.: Extracción de información con algoritmos de clasificación, Tesis, Instituto Nacional de Astrofísica Óptica y Electrónica, México (2005)