

Árboles de decisión para la evaluación del riesgo biológico de procesos biofarmacéuticos

Pavel, Novoa-Hernández^{1*}; Dailín, Cobos-Valdes²; Eduardo, Samaniego-Mena³; Milvio, Novoa-Pérez⁴

Resumen

En el presente trabajo se propone un nuevo modelo para la evaluación del riesgo biológico en procesos biofarmacéuticos. La propuesta extiende un modelo existente, aportando como principal novedad el tratamiento de las determinaciones de los niveles de consecuencia y probabilidad de riesgo, como problemas de clasificación supervisada. Específicamente, se obtuvieron modelos de clasificación basados en árboles de decisión que poseen como ventajas más importantes: 1) un número menor de indicadores para la determinación de consecuencias y probabilidades, 2) un orden de medición de los indicadores, basado en la importancia de los mismos. Con el objetivo de analizar las bondades del nuevo modelo, se consideraron tres casos de estudio relacionados con procesos farmacéuticos reales. En comparación con el modelo anterior, el nuevo ofrece resultados similares, pero facilitando notablemente el proceso de evaluación del riesgo biológico.

Palabras Clave: Árboles de decisión, Evaluación de riesgo biológico, Procesos biofarmacéuticos.

Decision trees for the biological risk assessment in biopharmaceutical processes

Abstract

A new model for assessing biological risk in biopharmaceutical process is proposed in the present work. This proposal extends an existing model including the handling of the consequence and probability levels computations as main novelty, and also as supervised classification problems. Specifically, two classification models based on decision trees were obtained, which gives as major advantages: 1) a lower number of indicators for the determination of consequence and probabilities, and 2) an order of measurement of the related indicators. In order to analyze the benefits of the new model, three real pharmaceutical processes were considered as cases studies. In comparison with the previous model, the new one offers similar results, but significantly facilitating the biological risk assessment process.

Keywords: Biological risk assessment, Biopharmaceutical processes, Decision trees.

Recibido: 05 de marzo de 2018
Aceptado: 25 de septiembre de 2018

¹Doctor dentro del programa de doctorado en Tecnologías de la Información y la Comunicación, Máster en Ciencias de la Computación e Inteligencia Artificial, Docente Titular (Profesor Agregado 1) de la Facultad de Ciencias de la Ingeniería de la Universidad Técnica Estatal de Quevedo, Ecuador; pnovoa@uteq.edu.ec; <https://orcid.org/0000-0003-3267-6753>

²Doctora en Ciencias de la Salud, Profesora Asistente; Production Laboratory Tech en Diagnostic Biochem, Canadá; dailin@cibho.hlg.sld.cu

³Master en Conectividad y Redes de Ordenadores; Docente contratado en Universidad Técnica Estatal de Quevedo, Ecuador; esamaniego@uteq.edu.ec

⁴ Doctor en Ciencias Filosóficas; Director General Adjunto para Posgrado e Investigación, Instituto Superior Politécnico Metropolitano De Angola, Angola; dapi@unimetroangola.co.ao

*Autor para correspondencia: pnovoa@uteq.edu.ec

I. INTRODUCCIÓN

En la actualidad, el empleo de agentes biológicos continúa siendo un tema de interés mundial debido al incremento de las enfermedades emergentes y reemergentes. De ello se deriva que todas las organizaciones que manipulen estos agentes, incluyendo materiales biológicos que pueden contenerlos, deben controlarlos de manera efectiva. En este escenario se encuentran representadas las entidades que ejecutan procesos biofarmacéuticos, en las que está presente el riesgo biológico.

La evaluación de este riesgo es un proceso de vital importancia para la toma de decisiones, la planificación y las actividades preventivas en materia de seguridad biológica. Existen diferentes métodos para evaluar el mismo que son seleccionadas teniendo en cuenta el objetivo que persigue la investigación y la experiencia de los investigadores en el tema, como por ejemplo (Jarne & Ferrarotti, 2003) y (Benavent, LLorca, Laborda, & Soto, 2004). Estos métodos en sentido general se realizan sobre la base de las políticas de seguridad y salud y seguridad biológica de cada país. En el contexto particular de Cuba, cada entidad con riesgos biológicos asociados aplica metodologías diferentes, pero no se establecen estándares para su ejecución.

Más recientemente, (Cobos, Vázquez, Cedeño, & Rodríguez, 2009) propone una metodología para la evaluación del riesgo biológico, sustentada en la base legal vigente en Cuba en materia de seguridad biológica, y ajustada a las características de cada entidad con este tipo de riesgo. No obstante de estos beneficios, dicha metodología puede ser objeto de mejora, especialmente en el análisis y comprensión de los resultados.

En este sentido, la presente investigación propone un nuevo enfoque para la evaluación del riesgo biológico basado en árboles de decisión. Los árboles de decisión son paradigmas computacionales que han sido ampliamente estudiados en los ámbitos de la Minería de Datos y Descubrimiento de Conocimiento (Ian H. Witten, Eibe Frank, Mark A. Hall, & Christopher J. Pal, 2017). En el caso particular de la investigación, se obtuvieron dos árboles de decisión a partir del algoritmo C4.5 (Quinlan, 1993), los cuales arrojaron como principales ventajas: 1) un número menor de indicadores para la determinación de consecuencias y probabilidades, 2) un orden de medición de los indicadores, basado en la importancia de los mismos. Con el objetivo de analizar las bondades del nuevo modelo, se consideraron tres casos de estudio relacionados con procesos biofarmacéuticos reales.

II. DESARROLLO

1. Trabajos relacionados

Actualmente, existen diversas metodologías para la evaluación del riesgo biológico que han sido aplicadas con éxito en diferentes escenarios. A continuación, se brinda una revisión de las metodologías más importantes que se relacionan con el tema de la presente investigación.

El método argentino BioRiM realizado en 2003 por (Jarne & Ferrarotti, 2003) se basa en las características intrínsecas que poseen los agentes biológicos, el tipo de proceso desarrollado, el área de trabajo, la duración y la frecuencia con que se llevan a cabo dichos procesos. Este método fue empleado por (Cedeño, 2010) para la evaluación del riesgo biológico en el proceso de diagnóstico de la tuberculosis humana con excelentes resultados.

Otro método empleado para la evaluación del riesgo biológico es Biogaval de procedencia española realizado en el 2004 por (Benavent et al., 2004). Su propósito es proporcionar al técnico de prevención un instrumento útil y práctico para valorar el riesgo biológico, así como para orientarlo en la priorización de las medidas preventivas y de control. Todo ello con una inversión de tiempo y dinero relativamente pequeña, especialmente cuando se tiene una cierta práctica en su manejo. Consta de los siguientes aspectos: determinación de los puestos a evaluar, identificación del agente biológico implicado, cuantificación de las variables determinantes del riesgo, medidas higiénicas adoptadas, cálculo del nivel de riesgo biológico e interpretación de los niveles de riesgo biológico. Este método quedó experimentalmente validado con un adecuado nivel de confianza, y fue aplicado por (Vázquez, 2011) para evaluar la exposición al riesgo biológico de los trabajadores del laboratorio clínico y microbiológico del hospital Dr. Gustavo Aldereguía Lima y sobre esta base se propuso un plan de acción. Los resultados fueron positivos en términos de prevención de riesgos.

Por su parte, (Cobos et al., 2009) desarrolló una metodología para la evaluación del riesgo biológico conformada por 9 pasos:

- 1) Identificación del proceso con riesgo y del material manipulado,
- 2) caracterización de algunos agentes orgánicos presentes en el material identificado,
- 3) clasificación del área según nivel de bioseguridad,
- 4) verificación del cumplimiento de los principios

- básicos de bioseguridad según el nivel de la misma,
- 5) identificación de los peligros biológicos,
 - 6) identificación de las causas específicas que provocan los peligros del paso 5),
 - 7) estimación de consecuencias,
 - 8) estimación de probabilidades, y
 - 9) determinación de la magnitud del riesgo.

La implementación de esta metodología para la evaluación del riesgo biológico se realizó mediante un procedimiento de normalización de operación (PNO) que integra seguridad y salud con bioseguridad para lograr una mejor gestión de los riesgos biológicos presentes en las entidades con este tipo de riesgo. Esta

metodología se ha aplicado en el Centro de Inmunología y Biopreparados desde el 2009 y en el 2012 se comienza a aplicar en el Centro para la Investigación y Rehabilitación de las Ataxias Hereditarias como parte de los resultados a generalizar de esta investigación.

Dentro de esta metodología, para la estimación de consecuencias y probabilidades se emplearon los aspectos que se observan en las tablas de los Anexos 1 y 2. Una vez determinado consecuencias y probabilidades, se estimó el riesgo empleando la tabla que se muestra en el Anexo 3. En sentido general estos pasos de resumen en la Figura 1.

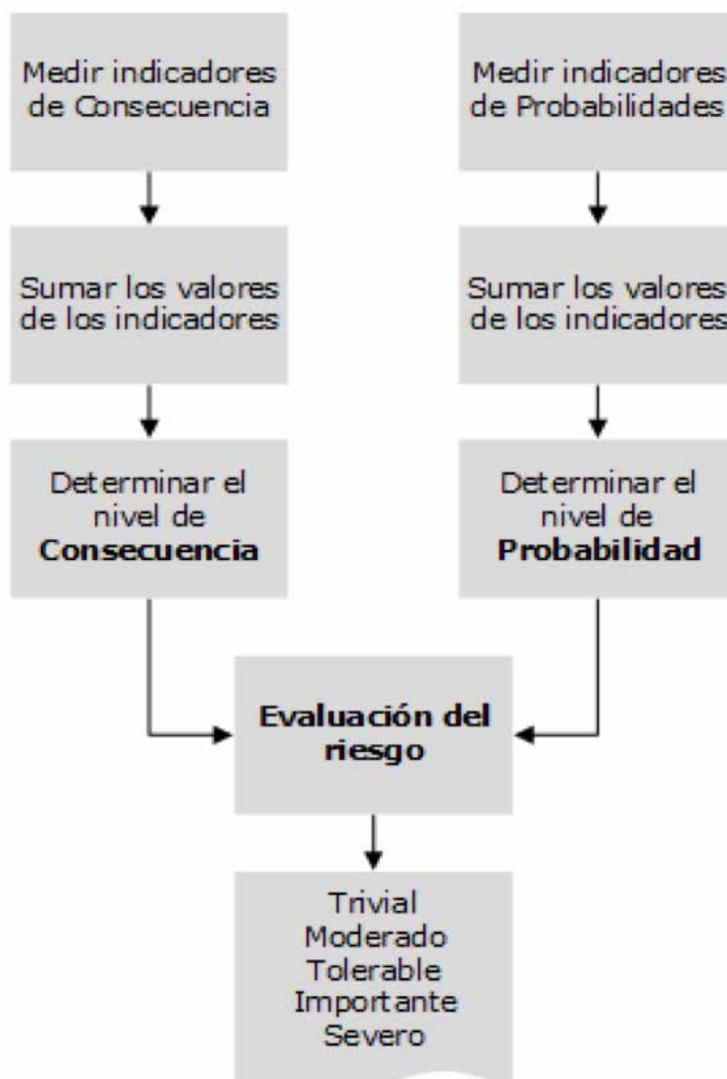


Figura 1. Metodología para la evaluación del riesgo biológico (Cobos et al., 2009)

La aplicación del aprendizaje automatizado en la evaluación de riesgos (Zio, 2018) es un tema aún en desarrollo. El aprendizaje automatizado o Machine Learning es una rama de la Inteligencia Artificial cuyo objetivo es desarrollar técnicas que permitan aprender a las computadoras. De forma más concreta, se trata de crear programas capaces de generalizar comportamientos a partir de una información no estructurada suministrada en forma de ejemplos (Shalev-Shwartz & Ben-David, 2014).

En el caso particular del ámbito médico y biofarmacéutico, sobresale la investigación de (Chazard, Preda, Merlin, Ficheur, & Beuscart, 2009), en la que se emplean árboles de decisión para identificar situaciones que puedan derivar en riesgo de reacciones adversas de medicamentos. En la obtención del modelo se tuvieron en cuenta 10500 registros de pacientes de Dinamarca y Francia, para generar 500 reglas.

En (Hathout & Metwally, 2016), los autores emplearon un algoritmo de aprendizaje automatizado para modelizar la carga de fármacos en nanopartículas de lípidos sólidos. Concretamente, aplicaron un proceso Gaussiano para correlacionar descriptores de los medicamentos.

La investigación desarrollada por (Reuelta-Zamorano et al., 2016) permitió obtener un modelo predictivo para identificar las infecciones asociadas a la asistencia sanitaria (HAIS) en una unidad de cuidados intensivos (UCI). Variables como la edad, la duración de la estadía, la cama donde se alojó el paciente y el mes de ingreso constituyeron los factores de riesgo más relevantes para predecir HAIS en la UCI que fue objeto de estudio.

Una revisión crítica sobre el desarrollo de modelos predictivos y dinámicos usando métodos de aprendizaje automático en la medicina fue llevada a cabo por (Alanazi, Abdullah, & Qureshi, 2017). Los autores reconocen el papel de los algoritmos de aprendizaje automatizado en diversos ámbitos de la medicina.

(Krämer, Schreyögg, & Busse, 2017) desarrollaron un modelo para clasificar las admisiones de pacientes hospitalizados según el diagnóstico primario: atención de emergencia o de atención programada. Particularmente, se predice el nivel de urgencia como un valor numérico. En el contexto de la oncología, (Lötsch et al., 2018) emplearon un algoritmo

clasificador de aprendizaje automatizado para predecir la ausencia o no, de dolor después de una cirugía de cáncer de mama.

El aprendizaje automatizado ha impactado otros escenarios de la sociedad moderna actual, algunos ejemplos son las investigaciones desarrolladas por (Ranginkaman, Kazemi Kordestani, Rezvanian, & Meybodi, 2014) y (Sikora, Krzystanek, Bojko, & Śpiechowicz, 2011). En la primera, se realiza un diagnóstico del drenaje en zonas costeras de Irán mediante el uso de métodos geoestadísticos, máquinas de soporte vectorial (SVM) (Campbell & Ying, 2011) en conjunto con un sistema de inferencia adaptativa neuro-difusa (ANFIS) (Abraham, 2005). En la segunda investigación, se aplica modelos de aprendizaje automatizado para describir y estimar en línea, el riesgo de contaminación por metano en minas.

1. Modelo propuesto

El modelo propuesto extiende al presentado por (Cobos et al., 2009), explicado en la sección anterior. Básicamente las principales diferencias radican en:

1. Uso de variables nominales, en lugar de variables numéricas, para definir los indicadores,
2. un árbol de decisión, para la determinación de los niveles de consecuencias y probabilidades, en lugar de la suma de los valores de los indicadores.

Para lograr la primera contribución se hizo corresponder a cada valor numérico del modelo de (Cobos et al., 2009), un valor nominal equivalente, como se muestra en la Tabla 1. El objetivo es facilitar el proceso de medición por parte de los especialistas. Nótese en este sentido que los valores nominales empleados resultan suficientemente intuitivos y por tanto, fáciles de establecer.

En relación a los árboles de decisión que determinan los niveles de consecuencia y probabilidad, se obtuvieron a partir de la aplicación de técnicas del Aprendizaje Automatizado. Específicamente se trataron ambos procesos, como problemas de Clasificación Supervisada, resueltos por el algoritmo C4.5 (Quinlan, 1993). El método utilizado es una implementación libre, incluida en la software Weka (Ian H. Witten et al., 2017).

Dado que, en ambos casos, se pueden generar todas las posibles combinaciones de valores para los indicadores, se procedió a utilizarlos como conjuntos

¹ En este tipo de problema, se genera un modelo de clasificación a partir de casos previamente clasificados.

de datos de entrada para el algoritmo C4.5. Como consecuencia, es de esperar que los árboles de decisión generados tengan una gran similitud con los resultados del modelo de (Cobos et al., 2009).

Tabla 1. Propuesta de correspondencia entre valores numéricos y valores nominales.

Factor a determinar	Indicador	Valor numérico según Cobos et al. (Cobos et al., 2009)	Valor nominal propuesto
Consecuencia	Grupo de riesgo	1	Grupo I
		2	Grupo II
		3	Grupo III
	Volumen de trabajo (Tamaño de la muestra)	1	Bajo
		2	Alto
	Daños al individuo	1	Bajo
		2	Moderado
		3	Alto
	Daños a la comunidad	1	Bajo
		2	Alto
	Tratamiento eficaz (¿Existe tratamiento?)	1	Sí
		2	No
	Medidas de emergencia (¿Plan de contingencia?)	1	SI
		2	No
Probabilidad	Valoración de los requisitos técnicos (¿Requisitos técnicos?).	1	Sí
		2	Parcial
		3	No
	Vías de infección y modo de transmisión. (Vías de infección)	1	Otras
		2	Pinchazos
		3	Aerosoles
	Cantidad de trabajadores expuestos	1	Pocos
		2	Muchos
	Niveles de exposición, frecuencia y duración	1	Bajo
		2	Apreciable
		3	Alto
		4	Excesivo
	Experiencia sobre ocurrencia de accidentes (¿Experiencia previa?)	1	SI
		2	No
	Niveles de inmunidad del personal (¿Inmunidad?)	1	SI
		2	No

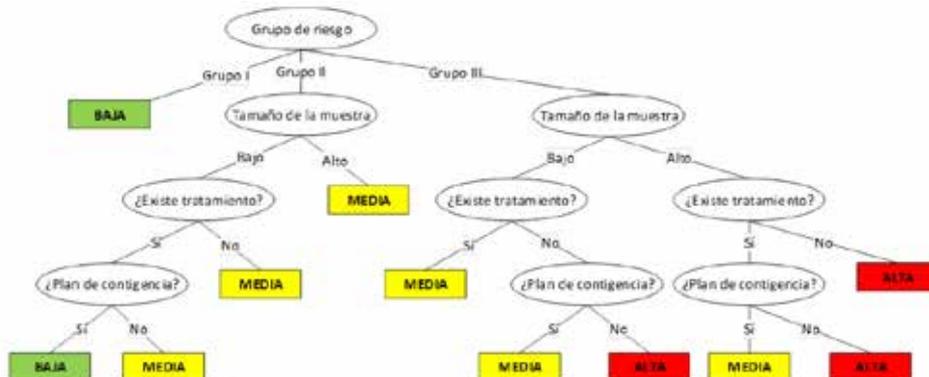


Figura 2. Árbol de decisión obtenido por el algoritmo C4.5 para la determinación del nivel de consecuencia.

En particular, el árbol correspondiente a la determinación de las consecuencias, mostrado por la Figura 2, clasifica los niveles de consecuencia de manera idéntica al modelo original. Sin embargo, como se aprecia en la propia Figura 2, este árbol ya no tiene en cuenta a los indicadores: Daños al individuo y Daños a la comunidad y el medio ambiente. Esto se debe a que estos indicadores están estrechamente relacionados con el indicador Grupo de riesgo, de manera que la información que pudieran aportar estos, resulta redundante para los

efectos de la determinación de la consecuencia. Por tal motivo, el algoritmo C4.5 no los considera relevantes. Asimismo, el árbol propone un orden para la obtención de los indicadores, y las ramas poseen longitudes diferentes. Esto implica que la determinación final del nivel de consecuencia no sigue los mismos pasos en todos los casos. Por ejemplo, cuando el Grupo de riesgo es el Grupo I, directamente el resultado es una consecuencia Baja, sin importar que valores tienen el resto de los indicadores.

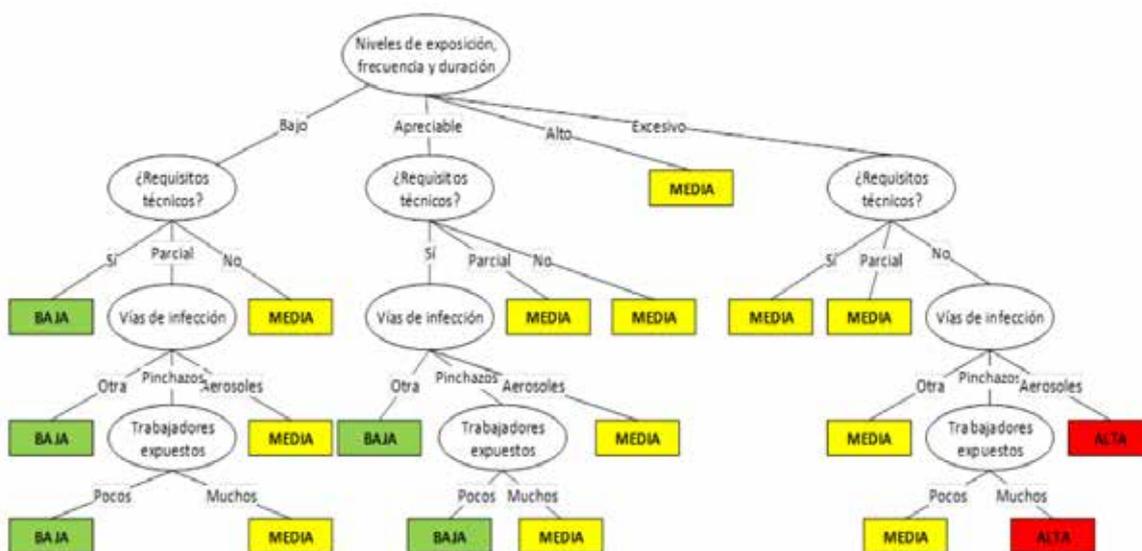


Figura 3. Árbol de decisión obtenido por el algoritmo C4.5 para la determinación del nivel de probabilidad.

Por otro lado, el árbol de decisión correspondiente a la determinación del nivel de probabilidad, se muestra en la Figura 3. En esta ocasión se alcanza un 83,3% de similitud con respecto al modelo original. Sin embargo, al igual que el caso anterior, el algoritmo C4.5 ha generado el árbol sin tener en cuenta determinados indicadores: Experiencia previa, e Inmunidad. Igualmente resulta interesante como en el caso de un valor Alto para el indicador Nivel de exposición, frecuencia y duración, se concluye directamente que la probabilidad es Media.

Finalmente, es importante mencionar que la manera en que el modelo propuesto evalúa al riesgo, es la misma que en el modelo original (Cobos et al., 2009) (véase Anexo 3).

2. Resultados y discusión

En este apartado se estudia la aplicabilidad del modelo propuesto. Primer lugar, se describen los casos de

estudios seleccionados, los cuales corresponden a tres procesos biofarmacéuticos reales. Más adelante, se realiza una comparación entre el modelo de (Cobos et al., 2009), y el propuesto.

Descripción de los casos de estudio

Los procesos seleccionados a continuación pertenecen al Centro de Inmunología y Biopreparados de Holguín (Cuba). En cada proceso existen riesgos asociados, los cuales se han identificado siguiendo la notación “CEN.n”, donde “n” es el número del proceso y “m” el del riesgo asociado.

1) Esterilidad de la Sangre bovina. Dedicado a la siembra de una muestra representativa de esta materia prima en medios de enriquecimiento, y Agar Triptona Soya para comprobar la esterilidad de la sangre. Este ensayo se realiza en un área controlada, donde el personal está en contacto con material biológico que puede estar contaminado.

El tiempo de incubación de las muestras varía en dependencia de la materia prima y el producto final que se esté ensayando. Los riesgos asociados son los siguientes:

- Contaminación por vertimiento de líquidos infecciosos (CE1.1).
- Contaminación por salpicaduras de líquidos infecciosos (CE1.2).
- Contaminación por cristalería rota contaminada (CE1.3).
- Contaminación por exposición a aerosoles (CE1.4).

2) Límite Microbiano. Ensayo microbiológico en áreas controladas, que se le realiza al Trofin-Vital como producto final. Este ensayo se realiza en un área controlada, donde el personal está en contacto con material biológico que puede estar contaminado. Con este análisis se obtiene una cuenta de bacterias y de hongos, así como la determinación de patógenos del producto final. Los riesgos asociados son los siguientes:

- Contaminación por salpicaduras de líquidos infecciosos (CE2.1).
- Contaminación por cristalería rota contaminada (CE2.2).
- Contaminación por exposición a aerosoles (CE2.3).

3) Producción de Trofin-Vital. Proceso tecnológico del Trofin-Vital. El grupo de trabajo está en contacto con material biológico, que ya se encuentra liberado por calidad. Después de conjunto de etapas (Hidrólisis y Mezcla final), se obtienen 527 frascos del producto mediante llenado manual. En este caso, los riesgos asociados presentes son:

- Contaminación por vertimiento de líquidos infecciosos (CE3.1).
- Contaminación por salpicaduras de líquidos infecciosos (CE3.2).
- Contaminación por cristalería rota contaminada (CE3.3).
- Contaminación por exposición a aerosoles (CE3.4).

4) Obtención de Sangre bovina. Se realiza en los mataderos del municipio donde el grupo de materias primas, plenamente capacitado, extrae generalmente de 4 a 6 botellones de sangre

bovina que conforman el lote. Esta operación se realiza tomando medidas asépticas para evitar desde la base, la contaminación del producto, que dentro de los requisitos de calidad tiene que ser una materia prima estéril para la producción de Trofin-Vital. Riesgos asociados:

- Contaminación por salpicaduras de líquidos infecciosos (CE4.1).
- Contaminación por pinchazos con agujas contaminadas (CE4.2).
- Contaminación por objetos corto punzantes contaminados (CE4.3).
- Contaminación por cristalería rota contaminada (CE4.4).
- Contaminación por exposición a aerosoles (CE4.5).

Los resultados de la aplicación del modelo propuesto y su comparación con el modelo de Cobos (Cobos et al., 2009), se aprecian en la Tabla 2. Se puede apreciar que el modelo generado automáticamente por el algoritmo C4.5 (J48) resulta muy similar al anterior. Específicamente, obsérvese que las predicciones realizadas por el modelo propuesto no subestiman a las del modelo anterior. Por ejemplo, en los riesgos C3.1-C3.3 y C4.1-C4.4, cuando el modelo de Cobos et al. obtiene un nivel de riesgo Trivial o Tolerable, el modelo propuesto predice Tolerable y Moderado respectivamente. En otras palabras, se ha obtenido un modelo muy similar al anterior, pero más conservador en cuanto a las predicciones que genera. Dado que el objetivo del modelo es contribuir a mitigar el riesgo biológico en procesos biofarmacéuticos, el hecho de que sea ligeramente conservador puede verse como un beneficio.

Sobre la posible difusión del modelo propuesto

Una interrogante que surge al crear un nuevo modelo que explique o ayude a decidir en un proceso como el que ocupa a la presente investigación es: ¿cómo aprovecharlo en escenarios reales? Téngase en cuenta que, aunque el modelo propuesto fue validado con el objetivo de compararlo con un modelo anterior, esto no significa que se pueda explotar eficientemente en entornos reales. Una posible respuesta a esta cuestión se puede encontrar en el ámbito de la automatización de procesos de gestión con el uso de las tecnologías de la información y la comunicación (TICs). Dado que el modelo obtenido tiene forma de árbol de decisión,

bastaría con incluirlo en un software que sea accesible por los especialistas que trabajan en procesos biofarmacéuticos con presencia de riesgo biológico. Lo anterior resultaría de gran utilidad práctica para dichos especialistas, a la vez que se logra difundir y aplicar los beneficios del modelo de manera efectiva.

Tecnológicamente hablando, en la actualidad existen varias opciones para conseguir dicho objetivo. Una de ellas es llevar el problema a RapidMiner

(Hofmann & Klinkenberg, 2013), a una herramienta similar a Weka, con el objetivo de exportar el modelo encontrado a un formato estándar de representación de modelo de datos. Precisamente, uno de los estándares más empleados es el denominado Predictive Model Markup Language (PMML) (Guazzelli, Lin, & Jena, 2012). Este estándar permitiría posteriormente la fácil inclusión del modelo en algún software que asista la toma de decisiones al especialista.

Tabla 2. Comparación entre los modelos de (Cobos et al., 2009), y el propuesto

Caso estudio	Modelo de Cobos et al.(Cobos et al., 2009)			Modelo propuesto		
	Consecuencia	Probabilidad	Riesgo	Consecuencia	Probabilidad	Riesgo
CE1.1	Media	Baja	Tolerable	Media	Baja	Tolerable
CE1.2	Media	Baja	Tolerable	Media	Baja	Tolerable
CE1.3	Media	Baja	Tolerable	Media	Baja	Tolerable
CE1.4	Media	Media	Moderado	Media	Media	Moderado
CE2.1	Baja	Baja	Trivial	Baja	Baja	Trivial
CE2.2	Baja	Baja	Trivial	Baja	Baja	Trivial
CE2.3	Baja	Media	Tolerable	Baja	Media	Tolerable
CE3.1	Baja	Baja	Trivial	Baja	Media	Tolerable
CE3.2	Baja	Baja	Trivial	Baja	Media	Tolerable
CE3.3	Baja	Baja	Trivial	Baja	Media	Tolerable
CE3.4	Baja	Media	Tolerable	Baja	Media	Tolerable
CE4.1	Media	Baja	Tolerable	Media	Media	Moderado
CE4.2	Media	Baja	Tolerable	Media	Media	Moderado
CE4.3	Media	Baja	Tolerable	Media	Media	Moderado
CE4.4	Media	Baja	Tolerable	Media	Media	Moderado
CE4.5	Media	Media	Moderado	Media	Media	Moderado

Fuente: Elaboración Propia

III. CONCLUSIONES

El modelo propuesto facilita el proceso de determinación de consecuencias y probabilidades, debido al empleo de valores nominales fáciles de establecer, y el uso de una menor cantidad de indicadores.

Aunque el modelo posee diferencias con el modelo previo (Cobos et al., 2009), las predicciones realizadas por éste, no subestiman a las del segundo; con lo cual, se logra un más modelo conservador. En la opinión de los autores de la presente investigación, esta característica es importante en modelos de decisión que predicen riesgos.

Las reglas generadas por los árboles de decisión obtenidos, pueden ser incluidas en un sistema de apoyo a la toma de decisiones, con el

objetivo de guiar al especialista en la evaluación del riesgo y que pueda contribuir a la protección del trabajador, la comunidad y el medio ambiente. Esto sería posible en un futuro cercano gracias a las tecnologías disponibles actualmente para el manejo y exportación de modelos de análisis de datos.

Dado el carácter general de los indicadores que definen al modelo, los resultados obtenidos en esta investigación pueden extenderse a otros escenarios que involucren riesgo biológico, particularmente en las instalaciones de salud que empleen un enfoque similar.

Como principales trabajos futuros se avizoran los siguientes: 1) incluir el modelo en una aplicación informática que esté más cercana al usuario final (especialista), y 2) optimizar el modelo aplicando

alguna técnica de computación inteligente, por ejemplo, con auto-adaptación (P. Novoa-Hernández et al., 2016; Pavel Novoa-Hernández, Corona, & Pelta, 2015).

IV. REFERENCIAS

- Abraham, A. (2005). Adaptation of Fuzzy Inference System Using Neural Learning BT - Fuzzy Systems Engineering: Theory and Practice. In N. Nedjah & L. de Macedo Mourelle (Eds.) (pp. 53–83). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/11339366_3
- Alanazi, H. O., Abdullah, A. H., & Qureshi, K. N. (2017). A Critical Review for Developing Accurate and Dynamic Predictive Models Using Machine Learning Methods in Medicine and Health Care. *Journal of Medical Systems*, 41(4), 69. <https://doi.org/10.1007/s10916-017-0715-6>
- Benavent, N., Llorca, J. L., Laborda, R., & Soto, P. (2004). *Manual práctico para la evaluación del riesgo biológico en actividades laborales diversas (BIOGAVAL) - VALENCIA*. Valencia.
- Campbell, C., & Ying, Y. (2011). Learning with Support Vector Machines. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 5(1), 1–95. <https://doi.org/10.2200/S00324ED1V01Y201102AIM010>
- Cedeño, M. (2010). *Evaluación del riesgo biológico en el diagnóstico bacteriológico de la tuberculosis humana mediante la aplicación del método Bioriesgo Intrínseco Mínimo (BioRIM)*. Instituto Superior de Tecnología y Ciencias Aplicadas, Ciudad de La Habana.
- Chazard, E., Preda, C., Merlin, B., Ficheur, G., & Beuscart, R. (2009). Data-Mining-based Detection of Adverse Drug Events. In A. H-P., B. B., J. Mantas, & I. Masic (Eds.), *Proceedings of MIE 2009. Medical Informatics in a United and Healthy Europe. Studies in Health Technology and Informatics* (pp. 552–556). Amsterdam: IOS Press.
- Cobos, D., Vázquez, J., Cedeño, M., & Rodríguez, A. (2009). Methodology to evaluate biological risks. *Revista Electrónica Ciencias Holguín*, XV(4).
- Guazzelli, A., Lin, W.-C., & Jena, T. (2012). *PMML in Action: Unleashing the Power of Open Standards for Data Mining and Predictive Analytics* (2nd ed.). Paramount, CA: CreateSpace.
- Hathout, R. M., & Metwally, A. A. (2016). Towards better modelling of drug-loading in solid lipid nanoparticles: Molecular dynamics, docking experiments and Gaussian Processes machine learning. *European Journal of Pharmaceutics and Biopharmaceutics*, 108, 262–268. <https://doi.org/10.1016/j.ejpb.2016.07.019>
- Hofmann, M., & Klinkenberg, R. (2013). *RapidMiner: Data Mining Use Cases and Business Analytics Applications*. Chapman and Hall/CRC.
- Ian H. Witten, Eibe Frank, Mark A. Hall, & Christopher J. Pal. (2017). *Data mining: Practical machine learning tools and techniques* (fourth edition) (2nd Editio). San Francisco: Morgan Kaufmann.
- Jarne, A. R., & Ferrarotti, N. F. (2003). Bioriesgo intrínseco mínimo: un método para la evaluación del riesgo causado por agentes biológicos. *Acta Bioquím. Clín. Latinoam.*, 37(1), 29–37.
- Krämer, J., Schreyögg, J., & Busse, R. (2017). Classification of hospital admissions into emergency and elective care: a machine learning approach. *Health Care Management Science*. <https://doi.org/10.1007/s10729-017-9423-5>
- Lötsch, J., Sipilä, R., Tasmuth, T., Kringel, D., Estlander, A.-M., Meretoja, T., ... Ultsch, A. (2018). Machine-learning-derived classifier predicts absence of persistent pain after breast cancer surgery with high accuracy. *Breast Cancer Research and Treatment*. <https://doi.org/10.1007/s10549-018-4841-8>
- Novoa-Hernández, P., Bayas, B. O., Oviedo, J. M., Puris, A., Menace, M., & Corona, C. C. (2016). Impacto de la auto-adaptación en ambientes dinámicos con frecuencia de cambio variable. *Investigacion Operacional*, 37(3).
- Novoa-Hernández, P., Corona, C. C., & Pelta, D. A. (2015). A software tool for assisting experimentation in dynamic environments. *Applied Computational Intelligence and Soft Computing*, 2015, 5.

- Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers.
- Ranginkaman, A. E., Kazemi Kordestani, J., Rezvanian, A., & Meybodi, M. R. (2014). A note on the paper “A multi-population harmony search algorithm with external archive for dynamic optimization problems” by Turkey and Abdullah. *Information Sciences*, 288, 12–14. <https://doi.org/10.1016/j.ins.2014.07.049>
- Revuelta-Zamorano, P., Sánchez, A., Rojo-Álvarez, J. L., Álvarez-Rodríguez, J., Ramos-López, J., & Soguero-Ruiz, C. (2016). Prediction of Healthcare Associated Infections in an Intensive Care Unit Using Machine Learning and Big Data Tools BT - XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016. In E. Kyriacou, S. Christofides, & C. S. Pattichis (Eds.) (pp. 840–845). Cham: Springer International Publishing.
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press.
- Sikora, M., Krzystanek, Z., Bojko, B., & Śpiechowicz, K. (2011). Application of a hybrid method of machine learning for description and on-line estimation of methane hazard in mine workings. *Journal of Mining Science*, 47(4), 493–505. <https://doi.org/10.1134/S1062739147040125>
- Vázquez, J. (2011). *Aplicación del Método BIOGAVAL para la evaluación del Riesgo biológico en el laboratorio Clínico y Microbiológico del Hospital Docente Clínico Quirúrgico “Dr. Gustavo Aldereguía Lima.”* Instituto Superior de Tecnología y Ciencias Aplicadas, Ciudad de La Habana.
- Zio, E. (2018). The future of risk assessment. *Reliability Engineering & System Safety*, 177, 176–190.