

UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS



**MODELO BASADO EN REDES BAYESIANAS PARA EL  
DIAGNÓSTICO DE LA FASCIOLOSIS BOVINA**

Tesis presentada en opción al Grado Científico de Doctor  
en Ciencias Técnicas

Neilys González Benítez

La Habana 2017  
“Año 59 de la Revolución”

UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS



**MODELO BASADO EN REDES BAYESIANAS PARA EL  
DIAGNÓSTICO DE LA FASCIOSIS BOVINA**

Tesis presentada en opción al Grado Científico de Doctor  
en Ciencias Técnicas

**Autor:** Ing. Neilys González Benítez

**Tutores:** Dra. C. Vivian Estrada Sentí

Dr. C. Antonio Romillo Tarke

Dra. C. Ailyn Febles Estrada

La Habana 2017

“Año 59 de la Revolución”

## *Agradecimientos*

- *A la persona que confió en mí, me apoyó y me aceptó; Ailyn Febles Estrada*
- *A mis tutores Vivian Estrada Sentí y Antonio Romillo Tarke, los que, con sus ideas, sabios consejos y críticas constructivas contribuyeron en gran medida a la realización de este trabajo y a mi formación científica.*
- *A Ramón Santana, por su valiosa en la revisión final del documento de tesis.*
- *Al PEFCEI, en especial a Febles, Rosa, Arturo, Alcides, Roxana, Olguita, Yamirys, Novo, Naivy y Lien, por el tiempo dedicado, la ayuda brindada y las valiosas contribuciones realizadas.*
- *A mis oponentes de predefensa de tesis, Dra. C. Arely Quintero Silveiro y Dr. C. Nemury Silega Hernández*
- *Al departamento provincial de Sanidad Animal, Empresa Agropecuaria Cubaquivir y Empresa Genética Camilo Cienfuegos de la provincia de Pinar del Río.*
- *Al Centro Meteorológico Provincial de Pinar del Río y al Instituto de Meteorología, por todo el apoyo.*
- *A todos mis amigos de verdad, especialmente a Máryuri y Hugo por apoyarme en los momentos más difíciles de mi vida.*
- *Y a todo el que, de una u otra forma contribuyó a que fuera posible llevar esta investigación a sus resultados finales.*

*Muchas gracias.*

## *Dedicatoria*

- *A mi hija Rocío, por ser la fuente de inspiración que me anima cada día*
- *A mi familia, en especial a mi madre y mi padre, por su apoyo constante*
- *A todos aquellos que me apoyaron en esta difícil tarea*
- *A la revolución cubana por permitirme estudiar, trabajar y por haber coincidido en una misma época de nuestra existencia*

## **SÍNTESIS**

La medicina enfrenta el reto de adquirir, analizar y aplicar conocimiento para resolver problemas clínicos complejos. En la actualidad existen innumerables adelantos que involucran el uso intensivo de la tecnología en el diagnóstico médico. La presente investigación está relacionada con el empleo de la tecnología en el diagnóstico de Fasciolosis bovina, por ser una de las enfermedades que más afecta al ganado bovino en Cuba ocasionando grandes pérdidas. El estudio realizado evidenció que las herramientas existentes para ejecutar diagnóstico de enfermedades en la ganadería, presentan características variables y no cumplen con los aspectos requeridos para el diagnóstico de la enfermedad objeto de estudio. En la presente investigación se propone un modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina. Se emplean las redes bayesianas por ser útiles para realizar predicciones en diferentes situaciones y calcular la probabilidad de una enfermedad concreta, con la presencia o no de algunos síntomas y otros datos relevantes. El modelo propuesto tiene tres componentes: gestión del conocimiento, análisis inteligente de los datos y las redes bayesianas. El aporte teórico de la investigación lo constituye la concepción y fundamentación de un modelo, basado en redes bayesianas, para el diagnóstico de la Fasciolosis bovina, capaz de brindar información con mayor grado de certeza, para apoyar la toma de decisiones. Con ello se logra un impacto en cuanto a la disminución de cantidad de animales enfermos y muertos. Para corroborar la validez del modelo, se aplicaron métodos cuantitativos y cualitativos.

## ÍNDICE

<b>INTRODUCCIÓN.....</b>	<b>1</b>
<b>CAPÍTULO I. MARCO TEÓRICO REFERENCIAL SOBRE LOS MODELOS PARA EL DIAGNÓSTICO DE ENFERMEDADES EN LA GANADERÍA .....</b>	<b>11</b>
<b>1.1.</b> La Fasciolosis bovina. Tratamiento de la información en la actividad veterinaria en Cuba .....	11
<b>1.2.</b> Inteligencia Artificial para asistir el diagnóstico de enfermedades en la ganadería.....	13
<b>1.3.</b> Técnicas de Inteligencia Artificial aplicadas a los diagnósticos de enfermedades.....	15
<b>1.3.1.</b> Propuesta de técnicas de Inteligencia Artificial para el diagnóstico de la Fasciolosis bovina.....	17
<b>1.4.</b> Redes bayesianas.....	21
<b>1.4.1.</b> Propiedades de las redes bayesianas.....	24
<b>1.4.2.</b> Tipos de redes bayesianas.....	26
<b>1.4.3.</b> Aprendizaje en redes bayesianas.....	28
<b>1.5.</b> Ventajas e inconvenientes de las redes bayesianas.....	34
<b>1.6.</b> Productos de software creado para implementar redes bayesianas útiles para el diagnóstico de enfermedades.....	36
<b>1.7.</b> Modelos para ejecutar diagnóstico de enfermedades en la ganadería.....	37
<b>1.8.</b> Aplicaciones de las redes bayesianas para uso veterinario.....	39
<b>1.9.</b> Conclusiones del capítulo.....	40
<b>CAPÍTULO II. MODELO BASADO EN REDES BAYESIANAS PARA EL DIAGNÓSTICO DE LA FASCIOSIS BOVINA.....</b>	<b>42</b>
<b>2.1.</b> Diagnóstico .....	42
<b>2.2.</b> Concepción metodológica del modelo.....	49
<b>2.3.</b> Estructura, componentes, cualidades, premisas y principios del <b>MRB-DiagPron</b> .....	50
<b>2.3.1.</b> Descripción general del modelo.....	53
<b>2.3.2.</b> Descripción de los componentes del modelo.....	54

2.4.	Conclusiones del capítulo.....	71
<b>CAPÍTULO III. VALIDACIÓN DEL MODELO PARA EL DIAGNÓSTICO DE LA FASCIOSIS BOVINA.....</b>		<b>72</b>
3.1.	Instanciación del modelo <b>MRB-DiagPron</b> .....	72
3.2.	Funcionamiento general de la aplicación.....	74
3.3.	Indicaciones metodológicas para la aplicación del modelo.....	75
3.4.	Diseño de la validación .....	77
3.5.	Valoración de los expertos sobre el modelo.....	77
3.6.	Desarrollo del experimento para el estudio del comportamiento de la Fasciolosis bovina.....	81
3.7.	Entrevista en profundidad.....	83
3.8.	Resultados de la triangulación metodológica de los métodos aplicados.....	84
3.9.	Comparación entre los modelos desarrollados para el análisis del comportamiento de la Fasciolosis bovina.....	85
3.10.	Comparación de la red bayesiana para el diagnóstico de la Fasciolosis bovina con las redes bayesianas existentes para uso veterinario.....	88
3.11.	Impacto del modelo.....	91
3.12.	Factibilidad económica.....	93
3.13.	Implicación social.....	97
3.14.	Conclusiones del capítulo.....	97
<b>CONCLUSIONES.....</b>		<b>99</b>
<b>RECOMENDACIONES.....</b>		<b>101</b>
<b>REFERENCIAS BIBLIOGRÁFICAS.....</b>		<b>102</b>
<b>ANEXOS.....</b>		<b>125</b>

## INTRODUCCIÓN

El diagnóstico médico es un ejemplo de toma de decisiones, ya que se requiere tener en cuenta varios aspectos como el historial clínico, la exploración física (o síntomas) y exploraciones complementarias (pruebas diagnósticas o clínicas) para determinar la presencia de la enfermedad más probable a padecer (Neapolitan, 2003).

La medicina (humana y animal) enfrenta el reto de adquirir, analizar y aplicar conocimiento para resolver problemas clínicos complejos. Se ha desarrollado en la actualidad innumerables adelantos que involucran el uso intensivo de alta tecnología como el diagnóstico por imágenes, avances quirúrgicos como la laparoscopia, robótica y farmacéuticos (Marasco y Doerfler, 2011). Monsalve (2017) refiere que se han realizado estudios apoyados en técnicas de Inteligencia Artificial (IA) que han demostrado efectividad en el área médica, los cuales han sido incorporados para apoyar la toma de decisiones y extraer conocimiento desde los datos. Según la referida autora, entre los estudios más destacados, apoyados en técnicas de IA, se encuentran los que emplean las redes bayesianas, las redes neuronales artificiales y los árboles de decisión, técnicas que han brindado resultados favorables en los procesos de diagnóstico y tratamiento de enfermedades.

En (Dietzel, Baltzer, Dietzel, *et al.*, 2011) y (Atkov, Gorokhova, Sboev, *et al.*, 2012) se realizan estudios médicos, haciendo uso de las redes neuronales artificiales, en (Chuang, 2011) y (Huang, Hung, Lee, Li y Wang, 2010); se presentan diagnósticos médicos haciendo uso del razonamiento basado en casos, y en (Tenorio, Hummel, Cohrs, Sdepanian, Pisa, *et al.*, 2011), se realizan estudios médicos apoyados en clasificadores bayesianos.

El desarrollo tecnológico actual ha permitido alcanzar grandes niveles en el almacenamiento de la información, mejorar con ello las posibilidades para desarrollar sistemas que permitan clasificar adecuadamente un diagnóstico, facilitar el diagnóstico diferencial y formular el



pronóstico y riesgo, basado en el procesamiento de la información.

En la actualidad la medicina utiliza adelantos que involucran el uso intensivo de alta tecnología, como el diagnóstico por imágenes, la robótica, entre otros, especialmente a través de la aplicación de métodos que aprovechan los datos disponibles y la experiencia clínica.

Para el cuidado y conservación de la salud animal, Cuba cuenta con herramientas para el procesamiento de la información respecto a las enfermedades más recurrentes en la ganadería, obteniéndose finalmente una estadística del comportamiento de dichas enfermedades.

A través del análisis documental y la experiencia personal de la autora, se detecta que la estadística que se obtiene está sustentada en el procesamiento de numerosos datos los cuales poseen problemas de almacenamiento, de calidad del dato, de gestión y análisis, por lo que se requiere de mecanismos y herramientas que contribuyan a mejorar estos procesos, para lograr un análisis más eficiente de la información y así poder extraer conocimiento útil desde ellos para apoyar la toma de decisiones.

Se identificaron estudios que describen la aplicación exitosa de herramientas para el tratamiento de la información, útiles para ejecutar diagnóstico de enfermedades en áreas de salud y otras para realizar predicciones en diferentes áreas. Entre las herramientas empleadas se encuentran las desarrolladas desde inicios de los años 60, ejemplo de ellas son; CASNET, MYCIN, PIP e Internist-I, herramientas que mostraron facilidad para realizar diagnóstico de varios tipos de enfermedades (Patel, Shortliffe, Stefanelli, *et al.*, 2008). CASNET (*Causal ASSociational Networks*), es una herramienta que se utilizó para construir sistemas expertos para el diagnóstico y tratamiento de enfermedades. MYCIN, se desarrolló a inicios de 1970 y fue diseñado para diagnosticar ciertas infecciones microbianas, además de recomendar el tratamiento con la asignación respectiva de medicamento necesario para tratar la infección. PIP, es una abreviación para el *Present Illness Program*, se desarrolló en 1970 para simular el comportamiento que debe seguir el especialista al momento de clasificar la historia médica de un paciente, diagnosticándolo

o identificándolo dentro de algún cuadro de las enfermedades renales conocidas. El Internist-I, se desarrolló a inicios de 1980, su función se centró en la investigación de métodos heurísticos para tareas de diagnóstico diferencial usadas para la toma de decisiones clínicas (fue usado en diagnóstico de medicina interna).

En los años 90, se reforzaron los estudios en las herramientas basadas en IA, utilizando el precepto de desarrollar herramientas basados en las necesidades actuales, ejemplo de estas herramientas fueron; ICHT, desarrollada para reducir la mortalidad de los niños sobre todo en las áreas rurales (Mahabala, 1992). HERMES (*Hepatology Rule-based Medical Expert System*), desarrollada para el pronóstico de las enfermedades crónicas más comunes (Bonfa, 1993). SETH, desarrollada para el tratamiento del envenenamiento por el uso de drogas (Darmoni, Massari, Droy, Thierry Blanc, Moiro y Leroy, 1994). PROVANES, desarrollada para el diagnóstico de los pacientes críticos en anestesiología (Passold, 1995).

En los años 2000, se desarrollan herramientas basadas en técnicas de IA para el diagnóstico de enfermedades, ejemplo de estas son; el Mediktor, desarrollada en el año 2015 por la empresa *Teckel Medical*, herramienta que constituye el primer evaluador de síntomas avanzado del mundo, capaz de reconocer lenguaje natural para que los usuarios expresen cómo se siente con sus palabras. La IA de Mediktor, junto con técnicas de gamificación, puede utilizarse en el campo de la formación para dinamizar la enseñanza (García, 2015). En el propio año se desarrolla la aplicación inteligente a partir del razonamiento basado en casos para el diagnóstico Síndrome de Asperger, aplicación de utilidad para la creación de sistema basado en casos (Ochoa, 2015). En el año 2016 se desarrollan otras herramientas basadas en técnicas de IA para el diagnóstico médico, ejemplo de ella es la desarrollada para el diagnóstico del alzhéimer, la cual mediante la fusión de imágenes funcionales y estructurales basado en el uso de la técnica de aprendizaje profundo –conocido también como *Deep Learning* (Ortiz *et al.*, 2015).

En el área ganadera se han desarrollado modelos para ejecutar predicciones respectivas a las producciones ganaderas, ejemplo de ellos son los modelos soportados en métodos estadísticos tradicionales; entre los que se encuentran el modelo MásLeche, desarrollado en Colombia (García *et al.*, 2011); el modelo de crecimiento de bovinos en pastoreo, desarrollado en Cuba (Díaz *et al.*, 2014); el modelo para el pronóstico de la producción de leche, mediante métodos ARIMA (Sánchez *et al.*, 2014). Daut, Lahodny, Peterson y Ivanek (2016), desarrollan el modelo informático para predecir la propagación de enfermedades debido al comercio de mascotas.

Jones (2016) desarrolla el modelo informático que predice brotes de enfermedades zoonóticas en base a los cambios en el clima, el crecimiento demográfico y el uso del suelo, su objetivo principal está en la mejora y comprensión de la propagación de enfermedades de los animales a las personas.

Las herramientas y modelos referidos, resultaron útiles, obteniéndose con ellos logros favorables, pero con la dificultad de que se diseñaron para ser aplicados a un solo dominio, lo que limita su reutilización.

Desde el punto de vista del paradigma estadístico, se ha utilizado el enfoque bayesiano ya que los métodos basados en la estadística bayesiana son especialmente útiles en los contextos de diagnósticos médicos, ellos permiten combinar la información que se obtiene de las observaciones, con información inicial proveniente de expertos o de datos históricos (Berger, 2006).

En la bibliografía consultada se identifica que el enfoque bayesiano hace algunos años tenía como principal inconveniente el cálculo de las distribuciones de probabilidades a posteriori, pero en la actualidad, los algoritmos computacionales basados en simulación han permitido superar ese obstáculo, en particular a través del uso de los métodos de Montecarlo, basados en Cadenas de Markov, los que constituyen herramientas que ayudan a resolver problemas complejos, (Gilks *et al.*, 1996).

Ejemplo de redes bayesianas desarrolladas en la medicina veterinaria para el diagnóstico de enfermedades son las construidas por (McKendrick *et al.*, 2000), para el diagnóstico de enfermedades tropicales bovinas, las construidas por (Otto y Kristensen, 2004), útiles para evaluar los riesgos de infección con micoplasma en cerdos, los trabajos desarrollado por Geenen y colaboradores, en el año 2011, donde utilizaron clasificadores bayesianos ingenuos (*naïve bayesian classifiers*) para el diagnóstico clínico de la fiebre porcina clásica, (Geenen, Gaag, Loeffen y Elbers, 2011), obteniéndose resultados significativos para distinguir poblaciones de animales infectados de las no infectados, comparable a un constructo conocido como regla diagnóstica. Las redes bayesianas construida para el análisis de datos complejos de animales (Lewis, *et al.*, 2011). En el año 2016, (Ordoñez, Guamán, Patiño y Torres, 2016), desarrollan redes bayesianas para el diagnóstico de enfermedades infecciosas virales en pequeñas especies, (canina y felina).

Las referidas redes bayesianas desarrolladas para la medicina veterinaria no son apropiadas para utilizarlas en los diagnósticos de diferentes enfermedades ganaderas ya que el tratamiento de los datos que con ellas se realiza no es el más adecuado y las explicaciones de los resultados que ofrecen son limitados, además no están diseñadas para ser reutilizadas en otros dominios diferentes para el que fueron concebidos en sus inicios.

Las técnicas de análisis y modelación de datos han ido creciendo progresivamente en los últimos años, proporcionando solución a diferentes problemas mediante métodos estadísticos, aprendizaje automático y a través del uso de las técnicas de IA, todo lo cual es útil para la extracción de conocimiento sobre asociaciones entre variables, patrones y correlaciones. El desarrollo de las Tecnologías de la Información y las Comunicaciones (TIC) brindan posibilidades para el análisis de grandes volúmenes de datos a gran escala, en particular haciendo uso de la Minería de datos.

Desde los años 2004, se han desarrollado los llamados sistemas de ayuda a la toma de decisiones médicas (que, por su sigla en inglés, se denominan, *Clinical Decisión Support Systems (CDSS)*)

son menos ambiciosos y más efectivo. Los CDSS son herramientas informáticas diseñadas para ayudar a la toma de decisiones clínicas, en las cuales las características individuales de los enfermos se utilizan para generar recomendaciones que son presentadas a los médicos para su consideración (Bright, Wong, Dhurjati, Bristow, Bastian, *et al.*, 2012).

El uso de las técnicas de IA y su aplicación en los sistemas de ayuda a la toma de decisiones, para el diagnóstico médico, tiene como propósito mejorar la calidad de los procesos médicos – asistenciales en la predicción y clasificación de enfermedades, con el fin de enriquecer el cuidado de los enfermos (Monsalve, 2017).

Una de las enfermedades ganaderas que mayor afectación ocasiona es la Fasciolosis, enfermedad parasitaria que afecta a gran cantidad de animales, es una zoonosis causada por el trematodo *Fasciola hepática*, la infección se adquiere debido a la ingesta de diversos vegetales acuáticos crudos, algunos terrestres, o agua contaminados con metacercarias. Se estima que existen al menos 2.4 millones de personas infectadas en 70 países y se considera que donde existen casos de las parasitosis en animales, también existen casos humanos (Zumaquero, Sarracent, Rojas y Martínez, 2013).

Cabeza (2010) refiere que en Cuba la Fasciolosis bovina presenta una elevada incidencia, donde alcanza nacionalmente casi el 30% de los bovinos sacrificados, destacándose la provincia de Pinar del Río, en las que la incidencia sobrepasa el 40 %.

Según datos reportados por el departamento de Sanidad Animal en la provincia de Pinar del Río, se constató una alta focalidad por Fasciolosis bovina, un elevado número de animales susceptibles, enfermos, muertos y sacrificados y un alto decomiso de hígado que provoca afectaciones económicas considerables.

El procesamiento de la información se basa en el control estadístico (manual) del comportamiento de dichas enfermedades, para lo cual se emplean los resultados de la vigilancia epizootiológica. Procesar los datos (dispersos, desordenados, individuales) obtenidos de la

población objeto de estudio durante el trabajo de campo, tiene como finalidad generar resultados (datos agrupados y ordenados), a partir de los cuales se realizará el análisis según los objetivos trazados.

Basado en el estudio realizado se logró identificar las principales deficiencias y limitaciones que se presentan en los departamentos de Sanidad Animal en la provincia de Pinar del Río, con respecto a la gestión de los datos y al procesamiento de la información relativa a las enfermedades ganaderas que frecuentemente afectan la ganadería. Se destacan como principales deficiencias y/o limitaciones las siguientes:

- La calidad del dato: los datos que se registran poseen insuficiencias en cuanto a su gestión, almacenamiento, tratamiento y análisis. Se tienen datos incompletos y ambiguos.
- Es poco efectivo el análisis de la información que se procesa, afectando la certeza de los diagnósticos y en consecuencia la toma de decisiones.
- Los estudios epizootiológicos que se realizan no satisfacen las demandas de los especialistas de salud animal, ya que la información disponible no siempre se almacena correctamente y generalmente no es confiable, lo que afecta la certeza del diagnóstico de enfermedades en la ganadería, el que con frecuencia no tiene buenos aciertos.
- Los métodos utilizados para el procesamiento de la información emplean estrategias poco eficientes que no producen información significativa. No se analizan las relaciones de dependencia entre las variables, en particular las variables del dominio.
- En el análisis de los datos no se hace uso de algoritmos especializados y se emplean cálculos estadísticos tradicionales.
- No se realiza inferencia probabilística de una manera eficiente.
- No se estima la probabilidad posterior de variables no conocidas en base a las variables conocidas.
- Los diagnósticos son demorados lo que provoca que no se tomen las decisiones en el

momento oportuno.

A partir del estudio realizado se identificó el siguiente **problema científico**:

¿Cómo diagnosticar la enfermedad de Fasciolosis bovina, con un mayor grado de certeza?

✓ A los efectos de la presente investigación un mayor grado de certeza se refiere a superar los diagnósticos anteriores respecto a:

- Cantidad de posibles animales enfermos por Fasciolosis bovina
- Cantidad de posibles muertes por Fasciolosis bovina

**Objeto de estudio:** Proceso de diagnóstico de enfermedades.

**Campo de acción:** Modelos para el diagnóstico de enfermedades en la ganadería bovina.

Se plantea como **objetivo general**: Desarrollar un modelo para el diagnóstico de la Fasciolosis bovina, basado en redes bayesianas, capaz de brindar información con mayor grado de certeza para apoyar la toma de decisiones.

Para el cumplimiento del objetivo general se definen los siguientes **objetivos específicos**:

1. Construir el marco teórico de la investigación relacionado con el diagnóstico de enfermedades en la ganadería, en particular de la Fasciolosis bovina.
2. Diagnosticar el estado que presenta el procesamiento de la información en el departamento provincial de Sanidad Animal en la provincia de Pinar del Río, relacionado con el diagnóstico de la Fasciolosis bovina.
3. Fundamentar un modelo para el diagnóstico de la Fasciolosis bovina, basado en redes bayesianas.
4. Elaborar una herramienta informática como soporte al modelo.
5. Validar el modelo propuesto a través de los métodos definidos en la investigación.

Según lo expuesto y derivado de la construcción del marco teórico referencial de esta investigación se plantea la siguiente **hipótesis de investigación**: Un modelo para el diagnóstico

de la Fasciolosis bovina, basado en redes bayesianas, que sea capaz de brindar información con mayor grado de certeza para apoyar la toma de decisiones.

Los métodos científicos que se utilizan para la presente investigación son: **Análisis-síntesis**, para el estudio de las fuentes bibliográficas existentes referente al tema, identificando los elementos más importantes y necesarios para dar solución al problema planteado. **Inducción-deducción**, para el estudio de las principales iniciativas de modelos para el diagnóstico de enfermedades, que permita reflejar los elementos comunes entre los fenómenos estudiados, establecer generalidades y analizar cada uno de los detalles hasta establecer las relaciones existentes. **Histórico-lógico y el dialéctico**, para el estudio crítico de los trabajos anteriores y utilizar estos como punto de referencia y comparación con los resultados alcanzados. **Modelación**, para la representación explícita de la solución propuesta a través de la modelación de los diagnósticos de enfermedades en la ganadería bovina, específicamente para la Fasciolosis bovina. **Sistémico estructural**, para que todos los elementos de la estructura y concepción del modelo funcionen como un todo. **Encuesta**, para obtener información sobre el tratamiento de la información que hace en los departamentos de vigilancia epizootiológica de las instituciones agropecuarias y la concepción de una herramienta para gestionar información, analizar y realizar diagnóstico de enfermedades en la ganadería bovina. **Análisis documental**, en la consulta de la literatura especializada en las temáticas afines de la investigación. **Experimento**, con el objetivo de evaluar en la práctica la validez del modelo en la ejecución de diagnóstico de enfermedades en la ganadería con mayor grado de certeza. **Entrevista a profundidad**, para obtener información referente a las enfermedades ganaderas, así como conocer los puntos de vista de especialistas con experiencias en salud animal. También se emplea en la validación de la propuesta. Se aplica la escala de *Osgood*, con el fin de validar los resultados del modelo propuesto para el diagnóstico de la Fasciolosis bovina. **Triangulación metodológica inter - método de forma simultánea**, con el objetivo de minimizar el sesgo de la investigación.



La **novedad científica** del trabajo se expresa en los siguientes aportes teóricos y prácticos.

**Aportes teóricos:** La concepción y fundamentación de un modelo, basado en redes bayesianas, para el diagnóstico de la Fasciolosis bovina, capaz de brindar información con mayor grado de certeza para apoyar la toma de decisiones.

**Aportes prácticos:**

- Una base de datos, que contiene información con mayor grado de confiabilidad, sobre las enfermedades en la ganadería bovina.
- Una herramienta informática, como soporte al modelo, basado en un sistema de información con las evidencias de los animales enfermos (síntomas y signos).
- Indicaciones metodológicas para la aplicación del modelo.

La tesis está estructurada de la siguiente forma: En el capítulo 1 se analizan los enfoques teóricos, con el objetivo de elaborar el marco teórico de la investigación relacionada con el diagnóstico de enfermedades, en particular en la ganadería. Se caracteriza de manera general la situación existente con el diagnóstico de la Fasciolosis. En el capítulo 2 se realiza un diagnóstico para conocer la situación que presentan las instituciones agropecuarias en la provincia de Pinar del Rio, la situación de la información que se gestiona y valorar la efectividad del análisis de la información que se procesa y su incidencia en la certeza de los diagnósticos. Se fundamenta y describe el modelo para el diagnóstico de la Fasciolosis bovina (**MRB- DiagPron**). Se presentan los principios, cualidades, premisas y componentes del modelo. En el capítulo 3 se presentan las indicaciones metodológicas para la aplicación del modelo y se valida la aplicabilidad del modelo y los principales resultados obtenidos.

Se presentan las **Conclusiones** y **Recomendaciones** derivadas de la investigación, las **Referencias bibliográficas**, así como los **Anexos**.

**Capítulo I: Marco teórico referencial  
sobre los modelos para el diagnóstico de  
enfermedades en la ganadería**

## **CAPÍTULO I. MARCO TEÓRICO REFERENCIAL SOBRE LOS MODELOS PARA EL DIAGNÓSTICO DE ENFERMEDADES EN LA GANADERÍA**

En el presente capítulo se abordan los elementos que conforman el marco teórico referencial y que componen el objeto de estudio de la investigación. Se realiza una caracterización de la Fasciolosis bovina. Se analizan las técnicas de Inteligencia Artificial aplicadas a los diagnósticos de enfermedades, relacionados con los modelos para el diagnóstico de enfermedades en la ganadería. Se provee un marco de referencia para interpretar las soluciones de los problemas resueltos haciendo uso de las redes bayesianas, en función del modelo a desarrollar, para el diagnóstico de la enfermedad objeto de estudio.

### **1.1. La Fasciolosis bovina. Tratamiento de la información en la actividad veterinaria en Cuba**

La Fasciolosis bovina es una infección parasitaria que produce pérdida de productividad en las producciones ganaderas, afecta a especies de animales tanto domésticos como silvestres y ocasionalmente al hombre. El agente productor de la Fasciolosis, se localiza en el hígado de los animales parasitados, la sintomatología y gravedad de esta parasitosis dependen de la intensidad de la infección y de las fases parasitarias que producen las lesiones (Zumaquero, Sarracent, Rojas, Martínez *et al.*, 2013).

En el ganado bovino existen cuatro infecciones parasitarias que tienen una elevada prevalencia en las explotaciones lecheras entre ellas se encuentran: *las protozoosis - cryptosporidiosis y giardiosis*; que afectan fundamentalmente a los animales en crecimiento, y *las helmintosis – gastroenteritis parasitarias y Fasciolosis, ligadas al pastoreo*; que afectan a animales de cualquier edad (Castro, González y Mezo, 2008).

En Cuba, la actividad veterinaria está regida por los departamentos provinciales de Sanidad

Animal, en ellos existe implementado el Sistema de Información y Vigilancia Epizootiológica (SIVE), sistema ordenado y continuo que contribuye al conocimiento de los cambios que se operan en las poblaciones animales en un espacio y tiempo determinado. Posee dentro de sus elementos: la notificación y registro de problemas zoonosarios, la investigación y registro de situaciones epidémicas, el estudio de casos individuales, así como la información sobre la infraestructura socioeconómica y la información sobre la población animal. Este sistema tiene limitaciones en cuanto al diagnóstico de enfermedades, en particular con la certeza de ellos, los cuales se realizan continuamente. Problemática que está dada por la desfavorable calidad que posee la información disponible, provocando incertidumbre para apoyar la toma de decisiones, la cual es requerida para cuidar la salud animal y conservar la masa ganadera.

Los datos que se registran en los departamentos de Sanidad Animal son utilizados para encontrar la prevalencia, morbilidad y mortalidad sobre cualquier enfermedad ganaderas de interés, resultados que se obtiene con técnicas estadísticas, en particular para los estudios relacionados con la búsqueda de factores de riesgo, donde se utiliza la estadística descriptiva y los modelos basados en regresiones para descubrir conocimiento sobre una población. Sin embargo, la búsqueda de relaciones relevantes entre los datos médicos no es buena ya que no se emplean técnicas que posibilitan encontrar nuevos conocimientos a partir de los datos que se tienen.

Por lo general en los departamentos de Sanidad Animal en Cuba, se maneja un elevado número de datos, relacionados con las informaciones relativas al comportamiento de la masa ganadera de forma general, y que son utilizados en este contexto para predecir el comportamiento de las enfermedades que con frecuencia afectan la ganadería cubana.

Tomando en consideración la importancia del significativo volumen de datos que se generan en los departamentos de Sanidad Animal, relacionados con síntomas clínicos de diversas

enfermedades, se requiere tratar los datos a través de procesos que no produzcan pérdida de información relevante y que puedan afectar de forma directa la calidad de los resultados obtenidos. En la presente investigación, se realiza un preprocesamiento de los datos con el fin de obtener datos de calidad, en aras de ejecutar el proceso de diagnóstico de enfermedades en la ganadería y obtener diagnósticos con mayor certeza para apoyar la toma de decisiones de los especialistas de salud animal.

El preprocesamiento de datos como paso previo tiene como objetivo, adecuar los datos, para que las técnicas aplicadas posteriormente puedan extraer la máxima información, y a su vez se pueda obtener resultados lo más certeros posible. En Bernal (2014) se reportan aplicaciones relacionadas con diagnósticos de enfermedades, donde previamente se ejecutó un proceso de Minería de datos, con el fin de obtener datos de calidad y desde los datos, ejecutar procesos de diagnósticos de enfermedades. Esas aplicaciones demostraron resultados favorables en diversas enfermedades de salud en humanos.

Es criterio de la autora de la presente investigación, que las aplicaciones reportadas en Calle (2014), en particular las relacionadas con la medicina veterinaria, no han sido las más adecuadas para ejecutar diagnósticos de enfermedades en la ganadería, debido a que no se tratan los datos previamente, siendo esta el área del conocimiento donde se generan diversos tipos de datos que presentan problemas de gestión y análisis, por lo que enfrenta el reto de encontrar nuevas soluciones, capaces de realizar procesamiento de la información para ejecutar procesos de diagnóstico de enfermedades con un mayor grado de certeza y apoyar la toma de decisiones.

## **1.2. Inteligencia Artificial para asistir el diagnóstico de enfermedades en la ganadería**

En la ganadería y en particular en la medicina veterinaria, existe la necesidad de encontrar nuevas soluciones para el procesamiento de la información, análisis desde los datos,

diagnósticos de enfermedades con un mayor grado de certeza. Posibilitando la obtención de resultados favorables para apoyar la toma de decisiones, cuidar la salud animal y conservar la masa ganadera.

Las técnicas de Inteligencia Artificial (IA), a partir de la década del 50, han sido utilizadas y aplicadas en diversas áreas comunes al desempeño cotidiano de las personas, ayudando, reemplazando, simulando las acciones o decisiones tomadas por individuos con ciertas características en particular. Como una de sus principales áreas de aplicación se puede destacar particularmente la medicina, dado que la Inteligencia Artificial logró su mayor impacto inicial en ella a través de los sistemas expertos y específicamente por medio de los sistemas de diagnóstico (Bernal, 2014).

La autora de la presente investigación, coincide con lo referido por el citado autor, en lo relativo al desarrollo de técnicas específicas pertenecientes a la Inteligencia Artificial, útiles para tratar diferentes problemas que requieren de mucho conocimiento, como es el caso particular del diagnóstico de enfermedades. Despierta un marcado interés y utilidad el empleo de las mismas, debido a las posibilidades que presentan para involucrarse en situaciones donde se requiere de un gran acervo de conocimientos, el veloz procesamiento de datos y la toma efectiva de decisiones.

Considera, además, la autora de la presente investigación, que los procesos de diagnóstico de enfermedades, y en particular los relacionados con la salud animal, serían más certeros al tratar los mismos con técnicas de Inteligencia Artificial. Evitando obstáculos que en ocasiones se presentan, debido a la coincidencia de uno o varios factores; como la escasez de expertos en el área médica o la abundancia/ausencia y dispersión de la información disponible. Influyendo, tales deficiencias sobre una variable crítica perteneciente al diagnóstico, el tiempo transcurrido desde

la observación de los síntomas, hasta la determinación de los agentes patógenos, responsables de la presencia de una determinada enfermedad.

Chávez, Silveira, Casas, Grau y Bello (2007) refieren, que las técnicas pertenecientes a la IA, se han utilizado con frecuencia para procesos de diagnósticos, por la facilidad que tienen para extraer información, en particular, a través de los sistemas basados en conocimiento, lo que es útil ya que permite realizar inferencia a partir del conocimiento a priori y llegar a una conclusión previa del problema que se analiza. Los elementos anteriormente presentados constituyen aspectos que justifican la posibilidad de aplicar técnicas de Inteligencia Artificial, para ejecutar diagnósticos de enfermedades en la ganadería.

### **1.3. Técnicas de Inteligencia Artificial aplicadas a los diagnósticos de enfermedades**

Las técnicas de Inteligencia Artificial aplicadas al diagnóstico de enfermedades han sido utilizadas en estudios de problemas complejos, alcanzando un aceptado grado de certeza en los resultados obtenidos con respecto a la identificación de un tipo específico de enfermedad. Estas aplicaciones son ventajosas debido a que facilitan la construcción y estudio de sistemas capaces de aprender a partir de un conjunto de datos y mejorar procesos de clasificación y predicción.

En Manickam y Abidi (1999), Alexopoulos, Dounias y Vemmos (1999), Zelic (1999), Ruseckaite (1999), se hace referencia a las aplicaciones de Inteligencia Artificial que fueron desarrolladas para ejecutar diagnósticos de enfermedades, útiles para ayudar a los médicos y pacientes a evitar diferentes complicaciones de enfermedades.

Lugo *et al.*, (2014) refiere que contar con ayuda computacional facilita el proceso de diagnóstico médico, porque lo hace sistemático y exhaustivo, no dependiente de la experiencia o del estado actual del médico. Elimina la necesidad de una memoria operacional que sobrepasa la capacidad humana.

Bernal (2014) cita que la Inteligencia Artificial es un área del conocimiento donde convergen diferentes técnicas, que tienen la finalidad de llevar a cabo tareas como el diagnóstico médico, estudios que han demostrado efectividad son:

- a) Las redes neuronales artificiales, (Dietzel, Baltzer, Dietzel, *et al.*, 2011), (Atkov, Gorokhova, Sboev, *et al.*, 2012).
  - b) El razonamiento basado en casos (Huang, Hung, Lee, Li y Wang, 2010), (Chuang, 2011).
  - c) Las redes bayesianas (Tenorio, Hummel, Cohrs, Sdepanian, Pisa, *et al.*, 2011).
- a) Las redes neuronales artificiales han demostrado que producen buenos resultados en bases de datos con muchas instancias. Las redes neuronales no consideran la incertidumbre, actúan como una caja negra, en el sentido de que no es posible saber cómo se ha llegado a los resultados obtenidos y los nodos intermedios no pueden ser interpretados (Rodríguez y Dolado, 2010).
- b) El razonamiento basado en casos se ha utilizado con éxito en el diagnóstico médico asistido, esta técnica se basa en el paradigma de resolución de problemas en el que un problema nuevo se soluciona atendiendo casos similares del pasado (Watson, 1999). El ejemplo más conocido, es el vecino más cercano, que consiste en comparar cada caso problema con una serie de casos almacenados; el caso problema nuevo y los casos almacenados consisten en un número de atributos o variables y éstos se comparan para encontrar el caso almacenado que más se parece al caso problema, de esta manera, se puede seleccionar la enfermedad conocida que más se parezca a la descripción del paciente. Una vez resuelto, este nuevo caso pasa a formar parte de la biblioteca de casos disponibles para comparar, con lo que el sistema se vuelve cada vez más sólido.
- c) Las redes bayesianas calculan la probabilidad de que un caso con cierto atributo pertenezca



a una categoría. En medicina se han usado para predecir recaída de enfermedad o riesgo de cáncer (Ramesh, Kambhampati, Monson y Drew, 2004), (Patel, Shortliffe, Stefanelli, Szolovits, *et al.*, 2009). El filtro no sabe esto por adelantado y tiene que aprender de la mano del usuario. El programa calcula en lo sucesivo la probabilidad de que cada atributo corresponda a una categoría o a otra.

En 2011, Geenen y colaboradores utilizaron clasificadores bayesianos ingenuos (*naïve bayesian classifiers*) para el diagnóstico clínico de la fiebre porcina clásica. Obtuvieron una exactitud de 60 a 70% para distinguir poblaciones de animales infectadas de las no infectadas, comparables a un constructo conocido como regla diagnóstica.

### **1.3.1. Propuesta de técnicas de Inteligencia Artificial para el diagnóstico de la Fasciolosis bovina**

En el presente epígrafe se analizan las técnicas de Inteligencia Artificial, utilizadas con frecuencia para ejecutar diagnóstico de enfermedades, para este análisis, la autora de la presente investigación utiliza la Teoría de Decisión Multicriterio Discreta (DMD), definida en Barba-Romero (1994) con el fin de seleccionar la técnica que mejor cumpla los criterios previstos para ejecutar el diagnóstico de enfermedades y en particular el de la Fasciolosis bovina. La teoría de la DMD, es considerada como una metodología útil para el apoyo a la toma de decisiones, ha resultado idónea en varios campos de aplicación, específicamente donde hay que decidir entre varias alternativas, teniendo en cuenta diversos criterios o puntos de vista.

Martínez y Escudey (1998) expresaron, que los métodos de evaluación y decisión multicriterio favorecen el proceso de encontrar soluciones posibles, pero no necesariamente óptimos, por lo que juega un importante papel en este proceso las preferencias del decisor y de los objetivos predefinidos.

# Modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina

## Capítulo I

Los citados autores, refieren que para un adecuado trabajo con los métodos multicriterio se debe tener en cuenta los siguientes aspectos:

1. Seleccionar la(s) mejor(es) alternativa(s).
2. Aceptar alternativas buenas y rechazar aquellas malas.
3. Generar una ordenación (*ranking*) de las alternativas consideradas (de la “mejor a la peor”).

Destacan que, para cumplir con los aspectos antes mencionados, existen diferentes enfoques, métodos y soluciones, a tener en cuenta, como son:

1. Ponderación Lineal (*Scoring*).
2. Utilidad Multiatributo (MUAT).
3. Relaciones de superación y procesos de Análisis Jerárquico (AHP – *The Analytic Process* – Proceso Analítico Jerárquico).

En la presente investigación, la autora, selecciona el método de ponderación lineal (*Scoring*), por el papel que juega el decisor para cada uno de los objetivos, con el fin de poder definir una estructura de preferencia entre las alternativas identificadas, tal como ocurre en la práctica en el caso del diagnóstico de enfermedades en la ganadería en las instituciones agropecuarias en Cuba. Su esquema básico consiste en construir una función de valor  $S_j$  para cada alternativa, de acuerdo a como se plantea en la expresión 1.1.

$$S_j = \sum_i w_i r_{ij} \quad (1.1)$$

Donde  $W_i$ ; es el peso del criterio  $j$  y  $r_{ij}$  la evaluación (*rating*) de la alternativa  $i$  respecto al criterio  $j$ .

A modo de guía para la selección de la técnica más adecuada para ejecutar diagnóstico de enfermedades, en Bernal (2014) se establece una correspondencia entre objetivos y técnicas de Inteligencia Artificial, empleadas en los diagnósticos de enfermedades más destacados y algunas de las herramientas desde la Inteligencia Artificial utilizadas en el área médica. Las definiciones formuladas en Bernal (2014) se ajustan a la presente investigación teniendo en cuenta la información extraída del análisis documental, en la tabla 1.1, se muestran los criterios y ponderaciones de las técnicas de Inteligencia Artificial que con frecuencia se han utilizado para el diagnóstico de enfermedades.

**Tabla 1.1.** Criterios y ponderación de las técnicas de Inteligencia Artificial a utilizar para el diagnóstico de enfermedades. Fuente: Elaboración propia.

<b>Criterios</b>	<b>Peso por criterio (<math>W_i</math>)</b>
1. Utilización de mecanismo de aprendizaje	5
2. Trabajo con datos que presentan incertidumbre	1
3. Posibilidad de combinar conocimiento y datos	3
4. Obtención de conocimiento verificable	5
5. Obtención de resultados de forma rápida y precisa	5
6. Grado de organización de la información	4
7. Utilización de analogías	3
8. Información excesiva o redundante	1
9. Flexibilidad	4

El resultado de la aplicación de la teoría de DMD aplicada a las técnicas de IA más apropiada para el diagnóstico de la Fasciolosis bovina se muestra en la tabla 1.2.

**Tabla 1.2.** Resultados de la Ponderación Lineal de cada criterio para la selección de la técnica de IA para el diagnóstico de enfermedades. Fuente: Elaboración propia.

<b>Criterios</b>	<b>Ponderación</b> $w_i$	<b>Razonamiento</b> <b>basado en casos</b> $r_{i_1}$	<b>Redes neuronales</b> <b>artificiales</b> $r_{i_2}$	<b>Redes</b> <b>bayesianas</b> $r_{i_3}$
Utilización de mecanismo de aprendizaje	5	7	9	9
Trabajo con los datos que presentan incertidumbre	1	3	7	9
Posibilidad de la combinación del conocimiento y los datos	3	8	6	9
Obtención de conocimiento verificable	5	8	1	9
Obtención de resultados de forma rápida y precisa	5	9	5	9
Grado de organización de la información	4	8	4	8
Utilización de analogías	3	9	1	2
Información excesiva o redundante	1	2	4	3
Flexibilidad	4	8	5	9
<b>Scoring <math>S_j</math></b>		<b>240</b>	<b>143</b>	<b>248</b>

Basado en el análisis realizado, la técnica de Inteligencia Artificial que reflejó resultados más apropiados, para el diagnóstico de enfermedades en la ganadería, se correspondió con el uso de las redes bayesianas. Resultado que está en correspondencia con el estado del arte relacionado con el uso de las técnicas de Inteligencia Artificial que con frecuencia se utilizan para el diagnóstico de enfermedades.

#### 1.4. Redes bayesianas

Las redes probabilistas son representaciones gráficas de las variables y de las relaciones entre las variables que caracterizan un problema (Wiltaker, 1990). Las redes bayesianas son un tipo muy popular de redes probabilísticas (Charles River, 2004), que proveen información sobre las relaciones de dependencia e independencia condicional existentes entre las variables. La inclusión de las relaciones de independencia en la propia estructura de la red, hace de las redes bayesianas una buena herramienta para representar conocimiento de forma compacta pues se reduce el número de parámetros necesarios. Estas relaciones simplifican la representación de la función de probabilidad conjunta como el producto de las funciones de probabilidad condicional de cada variable (Chávez, Grau y Casas, 2008).

Al representar una distribución de probabilidad, las redes bayesianas tienen una semántica clara, lo que permite procesarlas para hacer diagnóstico, aprendizaje, explicación, e inferencias (Heckerman, 1996). Según la interpretación, pueden representar causalidad y se refieren como redes causales (Spirtes, Glaymour y Sheines, 1993), (Pearl, 1993), pero no necesariamente tienen que representar relaciones de causalidad, sino de correlación (Grau *et al.*, 2004).

Según (Stuart y Norvig, 2003), una red bayesiana se define como un grafo que cumple lo siguiente:

- Los nodos de la red son variables aleatorias que se denotan con la letra X o con subíndices  $X_1, X_2, \dots, X_n$ . En principio estas variables pueden representar rasgos o atributos, pero puede ocurrir también que un rasgo original tenga que ser descompuesto en varias variables aleatorias. Por ejemplo, si el rasgo tiene múltiples valores puede desearse trabajar con variables aleatorias dicotómicas, una por cada valor del rasgo original.

- Cada par de nodos se conecta entre sí mediante arcos dirigidos. El significado de un enlace que va del nodo X al nodo Y es el de que X ejerce una influencia directa sobre Y. En términos de probabilidades esto significa que hay una dependencia condicional de Y respecto a X, esto es que la probabilidad de Y es diferente de la probabilidad de Y dado X.
- Por cada nodo hay una tabla de probabilidad condicional que sirve para cuantificar los efectos de los padres sobre el nodo. Los padres de un nodo son aquellos nodos cuyos arcos apuntan hacia éste.
- El grafo no tiene ciclos dirigidos (por lo tanto, es un DAG). Esto significa que no se presentan ambigüedades en el encadenamiento de probabilidades condicionales por el hecho de influencias directas cíclicas.

La autora de la presente investigación refiere que las redes bayesianas, son vistas como un grafo junto con las tablas de probabilidad condicional y pueden ser interpretadas como una representación aproximada de la función de distribución de probabilidad conjunta de la variable clase y de todos los rasgos predictores. La red en sí codifica un conjunto de pruebas de independencia condicional y las tablas de probabilidades condicionales completan la caracterización de la distribución conjunta. Todo lo cual contribuye a la realización del diagnóstico de enfermedades en la ganadería.

El grafo es importante para construir la red en sí, los valores que aparecen en las tablas de probabilidad condicional son imprescindibles en el procedimiento de inferencia (García, 1990), (Castillo *et al.*, 1997).

Formalmente esta representación de la distribución de probabilidad conjunta se define en un modelo de redes bayesianas, como un par (G, P), donde:

- G es un grafo dirigido acíclico (DAG)

- $P = \{p(X_1 \vee \tau_1), p(X_2 \vee \tau_2), \dots, p(X_n \vee \tau_n)\}$ ; es un conjunto de  $n$  distribuciones de probabilidad condicionales, una por cada variable
- $X_i$  (nodos del grafo)
- $\tau_i$  es el conjunto de padres del nodo  $X_i$  en  $G$

El conjunto  $P$  define la distribución de probabilidad conjunta asociada, como muestra la expresión:

$$P(X) = \prod_{i=1}^n P(X_i | t_i) \quad X = X_1, X_2, \dots, X_n \quad (1.2)$$

Las redes bayesianas han sido utilizadas para realizar tareas de clasificación, en particular tareas de predicción, aun cuando existe una sola variable contenida en bases de datos, cuya variable actúa como un clasificador y a su vez representa el problema que se desea predecir, mientras que todas las demás variables son los datos almacenados en la base de datos conformando un conjunto de datos, para tratarlos como casos en diferentes procesos.

Estudios realizados por Chávez, Casas, Moreira, Bello y Grau (2008), son ejemplos de lo antes referido, en particular en el estudio de la predicción de interacciones de proteínas; el cual tuvo como objetivo, balancear los casos clasificados como positivos, con los clasificados como negativos, otro ejemplo lo constituyó la predicción de sitios de *splicing*, donde el principal objetivo fue reducir los falsos positivos, por otra parte se utilizaron en la predicción de la hipertensión arterial, con el fin de reducir los casos falsos negativos.

El reverendo Thomas Bayes (1702-1761) desarrolló el Teorema que lleva su nombre, desde ese momento el teorema ha tenido un gran impacto en la inferencia estadística debido a que capacita a quien lo usa a inferir la probabilidad de una causa cuando el efecto de la misma es observado (Stone, 2013).

Existen aplicaciones en diferentes ramas de la ciencia, que emplean las redes bayesianas para

hacer inferencia y buscar soluciones probabilísticas acertadas, como lo es para los procesos de diagnóstico y en particular para los diagnósticos médicos, ejemplo de su uso son los reportados en los trabajos realizados por (Pourret y Onísco, 2008), (Højsgaard, Edwards, Lauritzen y Shahbaba, 2012), (Nagarajan, Scutari, Lébre y Stone, 2013).

#### 1.4.1. Propiedades de las redes bayesianas

Las redes bayesianas poseen un conjunto de propiedades útiles para el proceso de construcción, entre estas propiedades se encuentran la propiedad de Markov, la cual se cumplen para una red bayesiana formada por el par  $(G, P)$ , donde  $G$  es un grafo acíclico dirigido (DAG), en que cada nodo representa una variable aleatoria  $X_i$  de  $X = \{X_1, \dots, X_n\}$  y  $P$  una distribución de probabilidad conjunta sobre  $X$ , según (Korb y Nicholson, 2004); si y sólo si, cada nodo  $x_i$  es condicionalmente independiente de sus no descendientes,  $nd(x_i)$ , dado sus padres,  $pa(x_i)$ , como se muestran en la expresión 1.3.

$$P(x_i \vee pa(x_i), nd(x_i)) = P(x_i \vee pa(x_i)) \quad (1.3)$$

Las propiedades de Markov en las redes bayesianas, permiten una definición de redes bayesianas equivalente a la propiedad de factorización de la distribución de probabilidad conjunta, donde todo par  $(G, P)$  que cumple la propiedad de Markov constituye una red bayesianas y toda red bayesianas formada por el par  $(G, P)$ , cumple la propiedad de Markov, al considerar recíprocas estas características, entonces toda red bayesiana definida por el par  $(G, P)$ , cumple con las propiedades equivalentes a la factorización y la propiedad de Markov. En la presente investigación el trabajo con redes bayesiana se realiza partiendo de la propiedad equivalente a la factorización.

Es de destacar que las redes bayesianas, utilizan una serie de hipótesis que facilita el cálculo de los valores de las probabilidades, debido a la independencia condicional y/o separación



direccional (llamada *d-separation*), que introduce Pearl (1988), como alternativa a las propiedades de Markov en DAG. La separación direccional constituye una alternativa importante en el trabajo con las redes bayesianas, y como consecuencia se tiene que la probabilidad conjunta de una red bayesiana se puede obtener como el producto de las probabilidades de cada nodo condicionadas a sus padres, proceso que se conoce como el teorema de factorización de la probabilidad (Viviani, 2014).

La referida autora, cita que los modelos basados en redes bayesianas, poseen en común la asignación de probabilidad como medida de creencia de una hipótesis dada, por lo que la inferencia en estos modelos resulta importante, ya que la misma, constituye el proceso de introducir nuevas observaciones y calcular las nuevas probabilidades que tendrán el resto de las variables. La inferencia en los modelos basados en redes bayesianas, según refiere (Viviani, 2014) consiste en calcular las probabilidades a posteriori de un conjunto de variables después de obtener un conjunto de observaciones.

El uso de las redes bayesianas constituye una técnica ventajosa frente a los sistemas expertos clásicos basados en reglas cuando se utilizan para apoyar la toma de decisiones. En el contexto de la inferencia (Pearl, 2000) y (Castillo, Gutiérrez y Hadi, 1997) refieren que los modelos basados en redes bayesianas resultan ventajosos, debido a que es posible realizar inferencias bidireccionales; esto es, desde los efectos a las causas y desde las causas a los efectos y encontrar la mejor explicación para un conjunto de datos (Gámez y Puerta, 1998).

La ventaja más importante de una red bayesiana en el ámbito de la inferencia estadística se deriva de su habilidad para realizar cálculos locales (Pearl, 2001). Esta propiedad permite que la actualización de probabilidades se lleve a cabo de forma eficiente sin tener que calcular todas las posibles combinaciones entre los niveles de las variables (que supone un incremento exponencial

de los cálculos a medida que aumenta el número de variables del modelo el número de estados por variable) cuando se incorpora determinado conocimiento a la estructura gráfica. A los efectos de la presente investigación este hecho resulta importante ya que se emplea para actualizar o inferir la probabilidad de que una hipótesis pueda ser cierta en el diagnóstico de la enfermedad objeto de estudio.

Basado en el estudio realizado relativo a las redes bayesianas, la autora profundiza en otros aspectos propios de ellas, favorable a tener en cuenta para seleccionar del tipo de red bayesiana a utilizar para el diagnóstico de la Fasciolosis bovina. Entre los aspectos necesarios para la selección del tipo de red bayesiana a utilizar en la presente investigación se encuentran los citados por Castillo, Gutiérrez y Hadi (1997), relativos al tipo de variable a utilizar para el modelado con redes bayesianas y el tipo de distribución que se considera para cada variable.

#### **1.4.2. Tipos de redes bayesianas**

Existen tres tipos de redes bayesianas que son; las redes bayesianas discretas o multinomiales, normal o Gaussiana y mixtas, ellas se utilizan en dependencia del tipo de variable con la que se va a trabajar. Si las variables del problema son todas discretas, el modelo asociado es una red bayesiana discreta o red bayesiana multinomial. En este caso, se considera que todas las variables aleatorias de  $X = \{X_1, \dots, X_n\}$  son discretas, es decir, cada variable sólo puede tomar un número finito de posibles estados. La distribución de probabilidad de cada variable, condicionada a sus padres, es multinomial, por lo que dicha distribución queda especificada por tablas de probabilidades para las posibles combinaciones de estados entre las variables involucradas (Viviani, 2014).

La referida autora señala que las redes bayesianas construidas para representar diagnósticos médicos comienzan por considerar la presencia (o ausencia) de síntomas, que pueden causar una

determinada enfermedad asociado a signos. Estas dos variables, enfermedad y signos, pueden causar a su vez otro síntoma y la enfermedad puede producir, además, diferentes síntomas. En estos casos todas las variables consideradas son discretas dicotómicas que tienen como posible resultado Verdadero (V) en caso de presencia del síntoma o enfermedad, y Falso (F) en caso de ausencia, resultados que se obtienen al aplicar la factorización y calcular posteriormente la probabilidad conjunta.

Refiere, además, que, si las variables del problema siguen una distribución normal, la red es una red bayesiana normal o Gaussiana, por tanto, la relación entre ellas es lineal. Sin embargo, esta no es la única alternativa; por ejemplo, en el trabajo de Main y Navarro (2009) se propone un modelo más general, considerando la distribución conjunta que pertenece a la familia Potencial Exponencial Multivariante (MEP), introducida por Gómez *et al.*, (1998). Otra variedad a estos modelos se puede encontrar en el trabajo publicado por Driver y Morrell (1995) donde se propone aproximar las densidades condicionadas usando distribuciones de sumas Gaussianas ponderadas. En las redes bayesianas Gaussianas, la distribución de probabilidad conjunta asociada a las variables  $X = (X_1, X_2, \dots, X_n)$  tiene distribución Normal Multivariante con función de densidad conjunta. La inclusión de ambos tipos de variables en las redes bayesianas da lugar a las redes bayesianas mixtas. Viviani (2014) refiere que existen otro tipo de red bayesiana; reconocidas como red bayesiana Dinámica, con la característica de que están condicionadas por el tiempo.

Basado en el análisis antes realizado, el tipo de red bayesiana a utilizar para el diagnóstico de la Fasciolosis bovina, según criterio de la autora de la presente investigación, son las redes bayesianas discretas o multinomial, debido a que todas las variables del problema son discretas, cada variable sólo puede tomar un número finito de posibles estados. La distribución de probabilidad de cada variable, condicionada a sus padres, es multinomial, por lo que dicha

distribución queda especificada por tablas de probabilidades para las posibles combinaciones de estados entre las variables involucradas.

La red bayesiana, a construir para modelar la Fasciolosis bovina y sus correspondientes tablas de probabilidad condicionadas, comenzaran por considerar la presencia (o ausencia) de un síntoma, que puede ser causa de la enfermedad, esta variable, puede causar signos y estar asociada a factores de riesgo, por lo que todas las variables son consideradas discretas dicotómicas que tienen como posible resultado verdadero (V) en caso de la presencia del síntoma o enfermedad, y falso (F) en caso de ausencia.

Para la construcción de redes bayesianas se ha de tener en cuenta que existen dos formas fundamentales de construcción que son: construcción de forma automática y construcción de forma manual, aunque también se puede incluir como otro proceso de construcción de redes bayesianas la combinación de ambos tipos, según refiere (Calle, 2014).

El conjunto algoritmos empleados para la construcción de redes bayesianas de forma automática se reconoce en la bibliografía como algoritmos de aprendizaje, estos algoritmos poseen la facilidad de extraer toda la información necesaria para este tipo de proceso (Calle, 2014), (Viviani, 2014), (Soler, 2014).

### **1.4.3. Aprendizaje en redes bayesianas**

El aprendizaje en redes bayesianas, radica en definir la red probabilística a partir de datos almacenados en bases de datos en lugar de utilizar el conocimiento de los expertos. Este tipo de aprendizaje ofrece la posibilidad de inducir la estructura gráfica de la red a partir de los datos observados y definir las relaciones entre los nodos. Pearl (1988), expone formas de aprendizaje en redes bayesianas que son:

1. El aprendizaje paramétrico
2. El aprendizaje estructural

El aprendizaje paramétrico; *es* aquel que a partir de la estructura de la red facilita obtener probabilidades *a priori* de los nodos raíz y probabilidades condicionales de las demás variables requeridas a través del uso de bases de datos y con el aprendizaje estructural, obteniéndose una estructura a partir de los datos almacenados en la base de datos, con sus respectivas relaciones de dependencia e independencia entre las variables existentes. Es de destacar lo referido por Gámez y Puerta (1998), relacionado con los algoritmos que aprenden de la estructura de las redes bayesianas y que se engloban generalmente dentro de las siguientes categorías:

1. Algoritmos que se basan en procedimientos; los cuales son capaces de buscar la mejor estructura en el espacio de posibles soluciones, midiendo la calidad de cada red candidata mediante funciones de evaluación y se caracterizan por el tipo de función y por el procedimiento de búsqueda.
2. Algoritmos basados en detección de independencias; que son aquellos que toman como entrada el conjunto de relaciones de independencia condicional y generan la red que mejor representan estas relaciones.
3. Algoritmos híbridos; que son los que se basan en la combinación de ambas metodologías.

El proceso de aprendizaje paramétrico tiene como objetivo, encontrar los parámetros asociados a una estructura determinada de una red bayesiana. En la presente investigación se utiliza, para encontrar los parámetros de la estructura de la red bayesiana para el diagnóstico de la Fasciolosis bovina, contando con datos almacenados en base de datos lo cual facilita obtener de forma automática todos los parámetros necesarios.

El aprendizaje en redes bayesianas, es un requisito fundamental para ejecutar tareas desde los

datos, útiles para que el proceso que sobre ellos se realice fluya de forma adecuada por lo que se requiere disponer de base de datos con suficientes datos almacenados, donde estén especificados los valores de cada variable en cada uno de los casos para conseguir la estructura gráfica de la red a partir de los datos observados y definir las relaciones entre los nodos.

En la presente investigación, se crea una base de datos que contiene información confiable, sobre las enfermedades más recurrentes en la ganadería. Los datos contenidos en la base de datos y los seleccionados para el diagnóstico de la Fasciolosis bovina, conforman las variables del problema a modelar, lo que simplifica el proceso de obtención de las probabilidades, facilitando la obtención de la estructura gráfica de la red bayesiana.

En Viviani y Calle (2014) se hace referencia al método para estimar las probabilidades deseadas a partir de la frecuencia de los valores de los datos observados o contenidos en la base de datos, reconocido como estimador de máxima verosimilitud (ML, *Maximun Likelihood*); método empleado en la presente investigación para el proceso de diagnóstico de enfermedades en la ganadería, en particular para la Fasciolosis bovina, debido a que se cuenta con datos sobre la enfermedad objeto de estudio, lo que hace posible obtener la prevalencia de la enfermedad, analizar la frecuencia de la aparición de la mismas a partir de los síntomas, signos y factores de riesgos que se informan a los departamentos de vigilancia epizootiológicas de las instituciones agropecuarias.

Los referidos autores, señalan que la estimación de probabilidades deseadas a partir de la frecuencia de los valores de los datos, se relaciona con la instanciación de variables de entrada y propagación de sus efectos a través de la red bayesiana que se construye, proceso que se reconoce como el de propagación de probabilidades, relacionado con el razonamiento

probabilístico, el cual consiste en propagar los efectos de las evidencias a través de la red bayesiana para conocer las probabilidades a posteriori de las variables.

Para la ejecución de este proceso, existen diferentes tipos de algoritmo que pueden ser empleados en dependencia del tipo de grafo que se utiliza en la red bayesiana a construir para el problema en cuestión, los algoritmos son:

1. Algoritmos de propagación en árboles
2. Algoritmos de propagación en poliárboles
3. Algoritmos de propagación en redes multiconectadas

En la presente investigación, la autora, propone el uso de algoritmos de propagación en poliárboles, ya que el proceso de diagnóstico de una enfermedad, según Pourret y Onísco (2008), y con lo que coincide la autora de la presente investigación, no se basa solo en el diagnóstico desde los síntomas observados o en la conclusión que arroja una simple prueba diagnóstica, sino, que es el resultado de observar toda la evidencia, es decir, todos los síntomas que se presentan y las relaciones que estos tienen con el desarrollo de posibles enfermedades, para encontrar los valores más probables y ejecutar el proceso de diagnóstico médico.

La autora de la presente investigación refiere, que las redes bayesianas que con frecuencia se utilizan para modelar los diagnósticos de enfermedades, dadas evidencias comunes entre varias enfermedades, tienen un mayor grado de complejidad en cuanto al desarrollo del proceso de inferencia, para contrarrestar tal situación, utiliza los algoritmos de propagación en poliárboles. Por otra parte, el uso del método de agrupamiento a través de árboles de unión, es otro método que puede ser utilizado para atacar el problema del grado de complejidad en cuanto al desarrollo del proceso de inferencia (Barber 2014).

Cita la referida autora, que el método de agrupamiento a través de árboles de unión, consiste en transformar la estructura de la red bayesiana, obtenida al principio de la modelación del problema, para obtener un árbol mediante agrupación de nodos, usando la teoría de grafos, cuya estructura son redes multiconectadas, donde el aprendizaje, en redes multiconectadas resulta ser el más difícil de los tres. De manera análoga, sucede en cuanto a la complejidad en el proceso de propagación de probabilidades.

Señala, además, que en la actualidad se utilizan dos tipos de métodos para el aprendizaje con estructuras multiconectadas de redes bayesianas, estos métodos son:

1. Métodos basados en medidas de *Score* y *Search* (ajuste y búsqueda)
2. Métodos basados en pruebas de independencia

Sin embargo, existen alternativas relativas a los métodos antes mencionados las cuales hay que tenerlas presente cuando se requiere una optimización, estas alternativas consisten en proponer de nuevo el problema de aprendizaje, hasta encontrar la mejor estructura que produzca un rendimiento deseable, con el mínimo número de arcos, según refiere la citada autora.

En la presente investigación, la autora, utiliza el algoritmo de propagación para poliárboles, las relaciones entre las variables requieren establecer criterios cuantitativos para poder calcular la dependencia entre variables, siendo este criterio principal, ya que sirve de guía en la construcción de la topología de la red. Red que se nutre de los síntomas y signos que se observan en los animales enfermos, con el objetivo de conocer la enfermedad más probable que se presenta.

El algoritmo de propagación en poliárboles es introducido por Rebane y Pearl (2000) extendiendo el algoritmo propuesto por Chow y Liu para poliárboles, el cual parte del esqueleto (estructura sin direcciones), facilita determinar las direcciones de los arcos probando las relaciones de dependencia entre todas las tripletas de variables existentes en el esqueleto. Estas tripletas de



variables forman tres tipos de conexiones entre los arcos las cuales son:

1. Arcos convergentes.  $X \rightarrow Y \leftarrow Z$
2. Arcos divergentes.  $X \leftarrow Y \rightarrow Z$
3. Arcos secuenciales.  $X \rightarrow Y \rightarrow Z$

Gráficamente, estas conexiones son las que se muestran en el (Anexo 1).

Establecido el análisis de los algoritmos de propagación, en particular para poliárboles y estudiadas las conexiones entre los arcos, que son las que conforman grafos dirigido, en la presente investigación, la propagación se realiza utilizando poliárboles con conexión entre los arcos del tipo divergente, basado en la secuencia de pasos que se exponen en el (Anexo 2). La secuencia de pasos que se describe no garantiza obtener todas las direcciones; desde el punto de vista práctico, el principal problema es que generalmente no se obtiene independencia absoluta (información mutua cero), sin embargo, la independencia absoluta en estos casos se obtiene desde el conocimiento que se adquiere de la experiencia de los expertos.

El proceso de instanciación de variables de entrada y propagación de sus efectos a través de la red, se le llama propagación de probabilidades, que consiste en propagar los efectos de las evidencias a través de la red para conocer la probabilidad a posteriori de las variables (Viviani, 2014). De esta forma, tras conocerse los síntomas de una enfermedad, se obtiene la probabilidad a posteriori asociada a sufrir la enfermedad, dada la evidencia acerca de los síntomas. Siendo la probabilidad inicial o a priori, la probabilidad de presencia de la enfermedad, y la verosimilitud, la probabilidad de que los aquejados con una enfermedad, tengan los síntomas descritos.

En la presente investigación se tiene en cuenta la evidencia, cuando se observan los

animales enfermos con la probabilidad de tener alguna enfermedad, cuando es descrito los síntomas presentes en los enfermos, se tiene evidencia del problema que se presenta, lo que resulta útil dicha evidencia, debido a que la red bayesiana actúa como mecanismo de inferencia, actualizando la probabilidad conjunta, y por consiguiente, todas las probabilidades de la red bayesiana hasta obtener un resultado para el diagnóstico de la enfermedad en estudio.

Calle (2014), refiere que cuando se va a construir una red bayesiana, de la cual no se tiene evidencia acerca de ninguna de las variables del problema, dichas evidencias pueden ser obtenidas a través del cálculo de las probabilidades marginales para cada variable, obteniéndose así una primera información sobre la distribución marginal de cada variable del problema, información útil para incorporarla a la red bayesiana que se crea para el análisis del problema en cuestión.

### **1.5. Ventajas e inconvenientes de las redes bayesianas**

Las principales ventajas que poseen las redes bayesianas son:

- Tienen una interpretación sencilla, apoyada en métodos probabilísticos.
- Pueden ser utilizadas sin variar la red para: clasificación, predicción y exploración de las relaciones condicionales.
- Son robustas, en el sentido que pequeñas alteraciones en el modelo no afectan demasiado a la estructura de la red. Las actualizaciones producen modificaciones sencillas en los resultados.
- Se pueden incorporar los conocimientos o creencias previas de un experto a través de la distribución a priori. Esto permite combinar los datos estadísticos con los conocimientos previos de un modo muy práctico y efectivo.
- Cuando se trata de problemas con datos faltantes son mucho más efectivas que otras

técnicas.

- Están especialmente diseñadas para descubrir las relaciones entre variables que son causales y no poder ser observadas de otro modo, permitiendo hacer predicciones más precisas.
- Existen situaciones en las que conseguir suficientes datos es difícil. Partir de una red bayesiana basada en información inicial obtenida de expertos puede resultar eficaz y constituir un factor decisivo que proporcione ventajas sobre otros métodos.
- Las redes bayesianas no se ven afectadas por el sobreajuste de los datos.

Los elementos antes mencionados resultan favorables para la presente investigación ya que a través del uso de las redes bayesianas es posible realizar el proceso de diagnóstico de la Fasciolosis bovina, brindando explicaciones certeras desde una determinada observación o desde el criterio de los expertos.

Desde el punto de vista del análisis de datos, las redes bayesianas son una herramienta útil para el diagnóstico médico, debido a que no suponen un concluyente modelo subyacente, son fácilmente interpretables, adaptables y permiten la incorporación de conocimiento a priori de forma cualitativa.

A pesar del conjunto de ventajas que aportan las redes bayesianas en el contexto del diagnóstico de enfermedades, existen inconvenientes en su uso, algunos de estos inconvenientes son los citados por Druzdel (1993), que señala que el principal problema que se presenta al hacer uso de las redes bayesianas es en lo relativo, a la escasa relación que se da en algunas situaciones entre el razonamiento lógico humano y los métodos probabilísticos.

La desventaja antes mencionada no constituye un problema en la presente investigación, debido a que se trabaja con datos almacenados en la base de datos, los cuales son tratados con métodos

probabilísticos que facilitan su interpretación, además se trabaja con expertos que contribuyen a un adecuado razonamiento lógico humano, con el fin de que los resultados que se obtienen sean apropiados y entendibles al representar dichos resultados en redes bayesianas lo cual resulta favorable para el apoyo a la toma de decisiones.

### **1.6. Productos de software creados para implementar redes bayesianas útiles para el diagnóstico de enfermedades**

El uso de productos de software juega un papel importante a la hora de construir la estructura de la red bayesiana, puesto que permiten representar la red gráficamente, de tal forma que sea comprensible por los involucrados en el desarrollo de esta. Para este trabajo se han experimentado tres herramientas eficientes, al momento de construir redes bayesianas, las cuales son: **Elvira** (Lacave, Luque y Díez, 2007), **OpenMarkov** (Bermejo, Oliva, Díez, y Aries, s.f) y **UnBbayes** (Matsumoto, y otros, 2011). Estas herramientas cuentan con interfaz gráfica para diseñar la red y un modo inferencia que permite probar todos los posibles resultados de acuerdo a los síntomas clínicos que se presentan. Las herramientas **Elvira** y **OpenMarkov** están diseñadas para trabajar en ambientes de escritorio, lo que es un inconveniente para las exigencias actuales al momento del desarrollo de software, a diferencia de **UnBbayes** que es una herramienta completa, orientada a ambientes web y de escritorio.

En Murphy (2005) se hace una comparación detallada de productos de software en redes bayesianas, resumiéndose algunos en la presente investigación en el Anexo 3. En KDnuggets (2008) se hace una descripción de otros productos discretizados en software propietario y software libre, ver resumen en Anexo 4 y 5, respetivamente. En el trabajo de Doldán (2007) se comentan los productos de software del Anexo 3 y otras direcciones de Internet con herramientas basadas en el uso de redes bayesianas.

Basado en el estudio realizado sobre los productos de software creados para implementar redes bayesianas útiles para el diagnóstico de enfermedades en el contexto de la presente investigación, la autora propone crear una herramienta informática, como soporte al modelo basado en redes bayesianas, para el diagnóstico de la Fasciolosis bovina, la cual se orienta a ambientes web y de escritorio con la facilidad de integrarse con el sistema informático **UnBbayes**, el cual, está orientado a ambientes web y de escritorio, lo que la convierten en un instrumento eficiente para integrarla con sistemas informáticos de cualquier tipo, cuentan con interfaz gráfica para diseñar la red y un modo inferencia que permite probar todos los posibles resultados, de acuerdo a los signos clínicos, que presentan los animales enfermos facilita el análisis de la red recorriendo y activando los nodos para brindar información relativa al diagnóstico de enfermedades.

### **1.7. Modelos para ejecutar diagnóstico de enfermedades en la ganadería**

El desarrollo de la ganadería en Cuba ha tenido patrones políticos, económicos y sociales similares a lo largo de su diversa geografía. Para los ganaderos la tierra y el ganado representan un prestigio que a menudo triunfa sobre el valor de los retornos económicos directos.

Desde mediados de los años 70, se han hecho esfuerzos para modelar la producción ganadera, los modelos que se han desarrollado han sido utilizados para obtener asignación extrema de nutrientes en los rebaños, donde la calidad, cantidad y costos de nutrientes varían estacionalmente.

La autora de la presente investigación señala, que, conforme a las transformaciones que ocurren en la actividad ganadera, la aplicación de tecnologías de la información y de principios empresariales se vuelve cada vez más importante. Al mismo tiempo, a causa de la globalización y los avances tecnológicos, se ha comenzado a utilizar mejores técnicas de procesamiento de

datos, a través del desarrollo de diferentes modelos para ayudar a la predicción y rentabilidad de las operaciones que en la ganadería se realizan continuamente, ejemplos de estos modelos son:

- Modelo matemático para el diagnóstico y pronóstico de Fasciola Hepática en el ganado bovino, permite la interpretación de los datos, se apoya en las pérdidas en cuanto al decomiso de hígados, dando lugar a conocer el diagnóstico a priori de la enfermedad y su pronóstico a posteriori y con ellos poder realizar una intervención de forma temprana con el fin de evitar la propagación del parásito (González, 2014).
- Gestión organizacional para el desarrollo local a través de un modelado matemático aplicado a la Medicina Veterinaria, permite analizar el comportamiento de la Peste Porcina Clásica (PPC), brinda un análisis estadístico completo en un período de tiempo, por zonificación y relaciona las zonas con condiciones climáticas más propicias para el desarrollo de la enfermedad (González, Miranda y Malagón, 2014).
- Modelo estadístico de evaluación genética del ganado Holstein de Antioquia, permite predecir el valor genético del animal, tomar decisiones acertadas en el manejo, la nutrición, la sanidad, el mejoramiento genético y la reproducción de un hato, (Arias, Cardona, Herrera y Cerón, 2014).
- Modelo de macrófagos infectados por *Leishmania (V.) braziliensis*, desarrollada para identificar fenotipos de infección inducidos por silenciamiento genético con shRNAs, en macrófagos infectados por *L. Viannia braziliensis*, (Ovalle, 2015).
- Modelo informático que predice brotes de enfermedades zoonóticas en base a los cambios en el clima, el crecimiento demográfico y el uso del suelo, desarrollado para la mejora y comprensión de la propagación de enfermedades de los animales a las personas (Jones,

2016).

De manera general los modelos analizados para ejecutar diagnóstico de enfermedades en la ganadería, presentan características variables y no cumplen con los aspectos requeridos para el diagnóstico de enfermedades en la ganadería y en particular para la enfermedad objeto de estudio. Hecho que muestra la necesidad de una herramienta de apoyo para realizar diagnósticos de enfermedades en la ganadería, con el fin de obtener un mejor grado de acierto en los diagnósticos que se ejecutan y que contribuya apoyar la toma de decisiones, con el fin de disminuir las afectaciones en la ganadería, cuidar la salud animal y conservar la masa ganadera.

### **1.8. Aplicaciones de las redes bayesianas para uso veterinario**

Las redes bayesianas desarrolladas para uso veterinario e identificadas en la bibliografía, se basan en el análisis de los datos del comportamiento de las enfermedades que con frecuencia afectan la salud animal, estas redes bayesianas son complejas en lo referente al modelado que se hace con ellas debido a la naturaleza de los datos, ellas contribuyen de forma efectiva con el proceso de estudios epizootiológicos requeridos para un dominio específico, han ofrecido resultados favorables en la medicina veterinaria, en particular para conocer el comportamiento de la diarrea bovina a través de la red bayesiana desarrollada para el análisis de datos complejos de animales, para el análisis del comportamiento y la aparición de enfermedades tropicales bovinas, a través de la construcción de la red bayesiana que se desarrolló para este fin, la explicación que se obtiene con esta red bayesiana, es limitada, debido a que los resultados que muestran no son claros para los médicos veterinarios y especialistas de salud animal en general.

De manera general, los productos de software basados redes bayesianas útiles para el diagnóstico de enfermedades, los modelos para ejecutar diagnóstico de enfermedades en la

ganadería, y las aplicaciones desarrolladas en redes bayesianas para uso veterinario, identificadas en la bibliografía consultada, presentan características variables, lo que las limita en cuanto a su reutilización, no están aptas para ejecutar el proceso de diagnóstico de la Fasciolosis bovina. Hecho que muestra carencia de herramientas que brinden soporte a la toma de decisiones basadas en redes bayesianas para uso veterinario y en particular para el cuidado de la salud animal y conservación de la masa ganadera.

### **1.9. Conclusiones del capítulo**

- En el análisis documental se constata la existencia de modelos para el diagnóstico de enfermedades. Ellos son insuficientes en cuanto a la gestión y almacenamiento de la información para apoyar la toma de decisiones
- Se analizan las técnicas de Inteligencia Artificial que con frecuencia se han utilizado para el diagnóstico de enfermedades, a través de los resultados obtenidos al utilizar la teoría de la DMD para determinar la técnica a utilizar en la presente investigación, se obtiene que las redes bayesianas es la técnica de Inteligencia Artificial que mayores posibilidades tienen para el diagnóstico de enfermedades en la ganadería, debido a la capacidad de modelar los procesos del razonamiento médico combinando el conocimiento de los expertos con los datos médicos que ellos aportan y los datos clínicos existentes
- Se propone el tratamiento de los datos con el fin de obtener datos de calidad y poderlos utilizar en los diagnósticos de enfermedades, en aras de obtener un mayor acierto en los mismos para apoyar la toma de decisiones, cuidar la salud animal y conservar la masa ganadera. Su aplicación en el sector de la salud (humana y animal) constituye un enfoque recomendable, aportando excelentes resultados
- Se fundamenta el desarrollo de un modelo basado en redes bayesianas para el diagnóstico



de enfermedades en la ganadería y en particular para la Fasciolosis bovina, teniendo en cuenta habilidades técnicas y no técnicas, desempeño profesional y características de los especialistas de salud animal, que permita disponer de personal preparado para ejecutar diagnósticos de enfermedades en la ganadería desde los datos generados por el Sistema de Información y Vigilancia Epizootiológica y el Sistema Estadístico Veterinario.

**Capítulo II: Modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina**

**CAPÍTULO II. MODELO BASADO EN REDES BAYESIANAS PARA EL DIAGNÓSTICO DE LA FASCIOLOSIS BOVINA**

En el presente capítulo se realiza un diagnóstico inicial para determinar las dificultades existentes en cuanto a la información generada en los departamentos de vigilancia epizootiológica. Se propone un modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina, con mayor grado de certeza, compuesto por los componentes de gestión del conocimiento, análisis inteligente de los datos, construcción de redes bayesianas y diagnóstico de enfermedades en la ganadería. Se describe la estructura, cualidades, principios y componentes del modelo, así como la interacción entre cada uno de estos según las entradas y salidas propuestas.

**2.1. Diagnóstico**

Un diagnóstico es el proceso mediante el cual se lleva a cabo un análisis para recopilar información que ayude a determinar la situación actual de la organización y detectar sus áreas de mejoramiento. Mediante un diagnóstico se trata de focalizar y evaluar un conjunto de variables que juegan un importante papel en la comprensión, predicción y control del comportamiento de un fenómeno determinado (Shull, Singer y Sjoberg, 2008).

El desarrollo acelerado de la tecnología ha permitido el almacenamiento y tratamiento de grandes volúmenes de información compuesta de diferentes tipos de datos, los cuales no siempre son tan precisos y completos como se hace necesario. Con frecuencia se dispone de datos imperfectos ya que habitualmente se emplean cuantificadores imprecisos (ejemplo: elevado, bastante joven, muy mayor, muy pequeño, demasiado grande, etc.), en ocasiones se cometen errores en los instrumentos de medida utilizados al obtener las informaciones, se presentan ausencia datos, etc.

## Modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina

### Capítulo II

A pesar de que la imperfección está presente de forma directa en los datos, todavía el número de técnicas dentro de la disciplina del Análisis Inteligente de Datos que permitan el tratamiento de forma explícita de la información es bastante escaso. De esta forma, si las técnicas no son capaces de tratar con datos imperfectos, estos datos son transformados en precisos y durante este proceso de transformación es posible que se produzca una pérdida de información relevante que afecte de forma directa a la calidad de los resultados esperados (Martínez, 2014).

Los datos disponibles en Cuba sobre la ganadería bovina y en particular el diagnóstico de enfermedades no está ausente de los problemas antes descritos. Se han diseñado, elaborado e implementado diferentes mecanismos para el control y tratamiento de la información de enfermedades y en particular lo relacionado con el comportamiento y la aparición de enfermedades zoonóticas en la ganadería bovina, sin embargo, se evidencia en las investigaciones realizadas los limitados trabajos de investigación que estudien a profundidad estos temas en Cuba, lo cual limita la producción de materiales que traten su estado actual. Es por ello que la presente investigación se planteó la realización de un diagnóstico que permita profundizar en lo antes planteado.

En el estudio realizado para el diagnóstico participaron especialistas de salud animal que por su experiencia y conocimientos pueden desempeñarse como expertos en el diagnóstico de enfermedades en la ganadería. Dichos especialistas pertenecen al Departamento Provincial de Sanidad Animal, Empresa Pecuaria Punta de Palma, Empresa Agropecuaria Cubaquivir en la provincia de Pinar del Río, lo que representa un 80 % del total de centros agropecuarios de la provincia. No se tuvieron en cuenta centros de otras provincias, porque están en proceso de nueva estructuración, aunque participaron algunos representantes en diversas sesiones de trabajo.

**El diagnóstico estuvo dirigido a:**

- a) Analizar el estado que presenta el flujo informativo relacionado con la vigilancia epizootiológica sobre las enfermedades que ocasionan afectaciones en la ganadería y en particular el análisis de la cantidad de animales susceptibles, enfermos y muertos por Fasciolosis bovina de la masa ganaderas en la provincia de Pinar del Río, con el fin de realizar diagnóstico médico de dicha enfermedad desde los datos que se tienen sobre ella.
- b) Valorar la efectividad del análisis de la información que se procesa y su incidencia en la certeza de los diagnósticos.
- c) Valorar la situación del sistema de información, en particular las evidencias de animales enfermos (síntomas, signos y factores de riesgo).
- d) Identificar las principales dificultades que limitan la realización de diagnósticos certeros sobre la presencia de Fasciolosis bovina.

Estos aspectos fueron analizados aplicando diversos métodos y técnicas de investigación como son:

**1. Análisis documental:** Se analizó la información que se genera en los departamentos de vigilancia epizootiológica, de las instituciones de Sanidad Animal, para identificar el tipo de información y calidad de la misma, su registro y las posibilidades de acceso por los especialistas de salud animal en Cuba. Un resumen de los principales dificultades y limitaciones detectadas son las siguientes:

- ✓ Las herramientas que se emplean para el procesamiento de la información y ejecución de diagnóstico de enfermedades en la ganadería son muy rudimentarias, lo que incide en que los especialistas hagan poco uso de la información que se procesa relativa a

las enfermedades ganaderas, para realizar diagnósticos de enfermedades.

- ✓ No se realiza inferencia probabilística de una manera eficiente. No se estima la probabilidad posterior de variables no conocidas en base a las variables conocidas.
- ✓ No se emplean algoritmos especializados para el análisis de los datos y se emplean cálculos estadísticos tradicionales.
- ✓ Es limitada la socialización de la información y el conocimiento con otras instituciones que analizan enfermedades en la ganadería para realizar diagnóstico de enfermedades.
- ✓ La gestión, almacenamiento y análisis de la información para el diagnóstico de enfermedades en la ganadería en Cuba presenta limitaciones. El acceso a la información en ocasiones se dificulta. No siempre está disponible.

**2. Entrevista a profundidad:** Fue realizada a 15 especialistas de salud animal de una población constituida por 30 especialistas. La selección de la muestra fue intencionada teniendo en cuenta que era necesaria la opinión de especialistas en la vigilancia epizootiológica. Estos 15 especialistas incluyen una muestra representativa de cada institución, las interrogantes se encuentran en el Anexo 6. Los principales resultados obtenidos son:

- ✓ El 100 % de los entrevistados consideran necesario que las instituciones agropecuarias en Cuba cuenten con una base de datos con información mejor organizada y más confiable sobre las enfermedades en la ganadería bovina. Destacan que la calidad del dato es fundamental.
- ✓ El 93 % expresan que la investigación que se desarrolla es necesaria y pertinente.
- ✓ El 100 % consideran muy oportuno e importante la incorporación de las Tecnologías

de la Información y las Comunicaciones (TIC) de forma más efectiva para mejorar la efectividad de los diagnósticos.

- ✓ El 87 % considera que los métodos utilizados para el procesamiento de la información emplean estrategias poco eficientes que no producen información significativa. No se hacen análisis de las relaciones de dependencia entre las variables.
- ✓ El 93 % destaca que el Sistema de Información y Vigilancia Epizootiológica es un sistema ordenado que contribuye al conocimiento de los cambios que se operan en las poblaciones animales en un espacio y tiempo determinado, pero posee limitaciones en la calidad y tratamiento de la información disponible, lo que hace que no sean certeros los diagnósticos de enfermedades que se realizan para apoyar la toma de decisiones con el objetivo de conservar la masa ganadera y cuidar la salud animal. La identificación de relaciones relevantes entre los datos médicos no es buena debido a que no se emplean técnicas que posibilitan encontrar nuevos conocimientos a partir de los datos que se tienen.

3. **Encuestas:** Se aplicó la encuesta a 20 especialistas de salud animal de una población constituida por 30 especialistas. La selección de la muestra fue intencionada teniendo en cuenta que era necesaria la opinión de especialistas en la actividad pecuaria y que también tuvieran conocimientos sobre el tratamiento de los datos y el procesamiento de la información relacionada con la vigilancia epizootiológica. Estos 20 especialistas constituyen una muestra representativa de cada institución.

Se realizaron dos encuestas aplicando cuestionarios previamente elaborados con lo que se conoció la valoración de expertos con respecto a temas vinculados al objeto de estudio. Se diseñaron dos instrumentos (Anexos 7 y 8), los que posibilitaron profundizar en el

## Modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina

### Capítulo II

cumplimiento de las políticas de gestión de información establecidas por el Instituto de Medicina Veterinaria a través del Sistema de Información y Vigilancia epizootiológica, implementado en las instituciones provinciales de Sanidad Animal.

La primera encuesta presentada en el Anexo 7 estuvo destinada a especialistas de salud animal de los departamentos de las instituciones agropecuarias donde se gestionan información sobre el comportamiento de las enfermedades ganaderas que con frecuencia afectan la ganadería en la provincia de Pinar del Río y la segunda encuesta presentada en el Anexo 8 destinada a los especialistas en Medicina Veterinaria que poseen dominio informático en cuanto al tratamiento de la información relacionada con la vigilancia epizootiológica.

Los principales resultados obtenidos fueron:

- ✓ El 60 % de los encuestados manifestó no tener acceso o tener acceso limitado a la información sobre las enfermedades ganaderas que con frecuencia afectan la ganadería, el 35 % expresó poder utilizarla, y el 5 % de los encuestados no respondió a la pregunta.
- ✓ En cuanto a la frecuencia con que participan los especialistas de salud animal en la elaboración conjunta de diagnóstico de enfermedades en la ganadería bovina; el 15 % de los encuestados expresó que nunca participan, el 40 % afirmó que pocas veces participan y el 45 % declaró que a veces participan, lo que demuestra que en las instituciones agropecuarias aproximadamente el 50% de los especialistas de salud animal no participan en la elaboración de diagnósticos de enfermedades, haciendo uso de la información disponible.
- ✓ En cuanto a la frecuencia con la que se utiliza la información que se procesa sobre las enfermedades ganaderas más recurrentes para la elaboración de diagnóstico de dichas



## Modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina

### Capítulo II

enfermedades, el 45 % de los encuestados respondió que nunca utilizan la información y el 20 % respondió que pocas veces la utilizan, el 25 % respondió que a veces utilizan la información y el 10 % no respondió la pregunta.

- ✓ Solo el 25 % de los encuestados hace uso de bases de datos con información generada desde el Sistema de Información y Vigilancia Epizootiológica, lo que demuestra el poco uso o desconocimiento de estas herramientas con la que se gestiona información en las instituciones de salud animal.
- ✓ El 100 % de los encuestados, expresó la necesidad de la creación de herramientas para diagnosticar la enfermedad de Fasciolosis bovina, con un mayor grado de certeza.
- ✓ El 100 %, estuvo de acuerdo en que la calidad de la información relativa a las enfermedades ganaderas puede mejorar, con el objetivo de realizar mejores diagnósticos de enfermedades.
- ✓ El 95 % planteó que los métodos utilizados para el procesamiento de la información emplean estrategias poco eficientes que no producen información significativa.
- ✓ El 90 % de los encuestados, destacan la necesidad de lograr una adecuada gestión y almacenamiento de la información, para garantizar una mejor calidad en los diagnósticos.

El análisis realizado en los departamentos de las instituciones agropecuarias, donde se gestiona información sobre el comportamiento de las enfermedades ganaderas que con frecuencia afectan la ganadería se evidenció:

- ✓ El 70 % de los encuestados plantea que los datos históricos no se almacenan de forma adecuada y estos son incompletos y presentan incertidumbre.
- ✓ El 60 % de los encuestados destacan que no existe un desarrollo de herramientas

informáticas que permita un mejor tratamiento de la información.

- ✓ Se destaca la necesidad de disponer de bases de datos, en las instituciones agropecuarias cubanas, para gestionar información sobre las enfermedades ganaderas que frecuentemente afectan la ganadería.

Esto evidenció la pertinencia de la investigación en cuanto a la necesidad de garantizar una base de datos normalizada con la información relativa a las enfermedades ganaderas que con frecuencia afectan la ganadería bovina para ejecutar diagnóstico de enfermedades, útil para apoyar la toma de decisiones, con el fin de cuidar la salud animal y conservar la masa ganadera. El resultado de este diagnóstico refleja la importancia que reviste disponer de información organizada de forma tal que facilite su almacenamiento, acceso y análisis, así como poder contar con tecnología apropiada para diagnosticar la Fasciolosis bovina con mayor grado de certeza y con ello apoyar la toma de decisiones.

## **2.2. Concepción metodológica del modelo**

Para la concepción metodológica del modelo se tuvieron en cuenta los resultados arrojados en el marco teórico referencial de la investigación, así como las consideraciones expuestas en el epígrafe 2.1 apoyadas en la realización de un diagnóstico preliminar a través de entrevistas a profundidad, encuestas y el análisis documental como métodos teóricos de la investigación. Los aspectos considerados, de acuerdo al análisis previo, fueron:

1. Cumplimiento del proceso de gestión de la información y vigilancia epizootiológica establecido por el Instituto Nacional de Medicina Veterinaria.
2. Estado de la gestión de la información relativa a las enfermedades ganaderas que con frecuencia afectan la ganadería bovina para ejecutar diagnóstico de enfermedades desde dicha información.

3. Necesidad de las instituciones agropecuarias de adoptar un modelo de diagnóstico de enfermedades en la ganadería, ajustable según las características de la tecnología.

En la bibliografía analizada se expone un grupo de investigaciones relacionadas con la aplicación de la tecnología en el diagnóstico de enfermedades, pero el objeto de análisis en ellos es diferente al de la presente investigación y por ello no satisface las necesidades del problema de investigación abordado por la autora. Por otro lado, la información de entrada es diferente en estructura, contenido, así como la fuente de procedencia. Estas investigaciones no integran herramientas para el tratamiento de información incompleta y con incertidumbre. Estos elementos hicieron necesario el estudio de la teoría y una evaluación de la situación del país, en particular la provincia de Pinar del Río, para elaborar un modelo que mejore los diagnósticos de la Fasciolosis bovina.

Partiendo del análisis realizado y las variables definidas en la presente investigación, se propone un modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina (**MRB-DiagPron**) que integre los procesos: gestión del conocimiento, análisis inteligente de los datos, construcción de redes bayesianas y diagnóstico de enfermedades en la ganadería.

### **2.3. Estructura, cualidades, premisas, principios y componentes del MRB-DiagPron**

La estructura, cualidades, premisas y principios del modelo **MRB-DiagPron** desarrollado en la presente investigación, se representa en la figura 2.1. El modelo **MRB –DiagPron** tiene una estructura abierta debido a que establece intercambio con el entorno general y específico. En el caso del entorno específico permite el intercambio con especialistas en salud animal. La interacción con el entorno general se establece de forma indirecta a través de: factores tecnológicos, ambientales y políticos-legales.

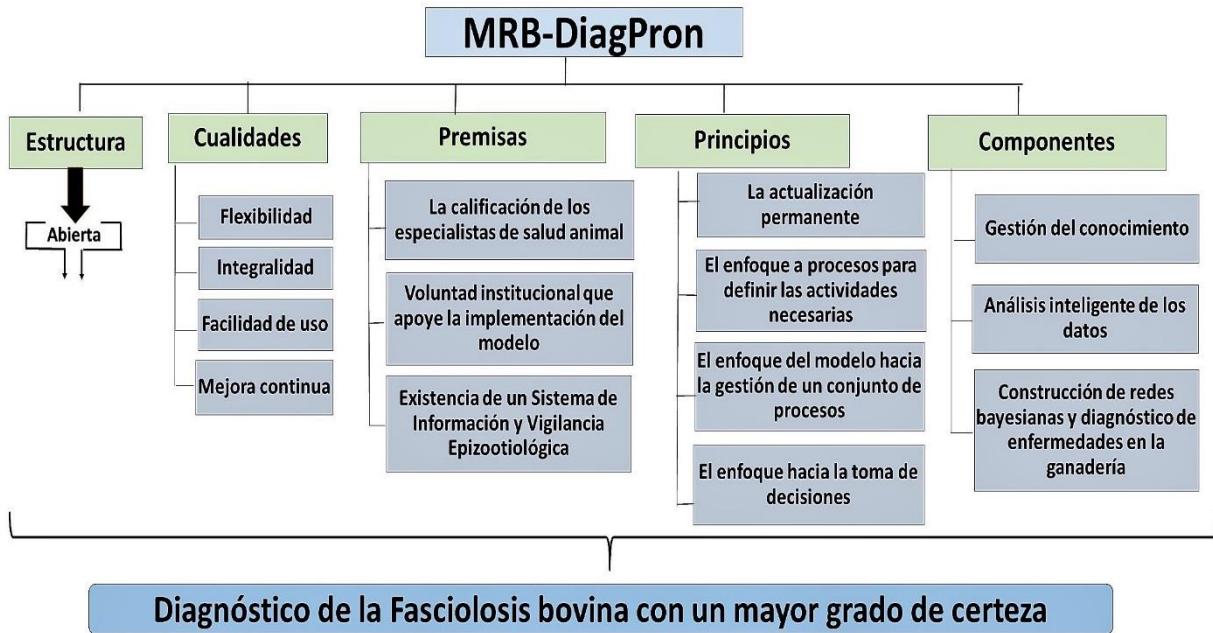


Figura 2.1. Estructura, componentes, premisas, cualidades y principios del MRB-DiagPron.

Fuente: Elaboración propia.

**Las cualidades que lo distinguen son:**

- ✓ Flexibilidad
- ✓ Integralidad
- ✓ Facilidad de uso
- ✓ Mejora continua

**Las premisas con vistas a la aplicación del modelo propuesto son:**

1. La **existencia** de un Sistema de Información y Vigilancia Epizootiológica útil para conocer el comportamiento de la salud animal en las producciones agropecuarias.
2. La **calificación** de los especialistas de salud animal necesaria para el uso eficiente del modelo propuesto para el diagnóstico de enfermedades en la ganadería.

3. La **voluntad** institucional para apoyar la aplicación del modelo.

**El modelo se sustenta bajo los principios de:**

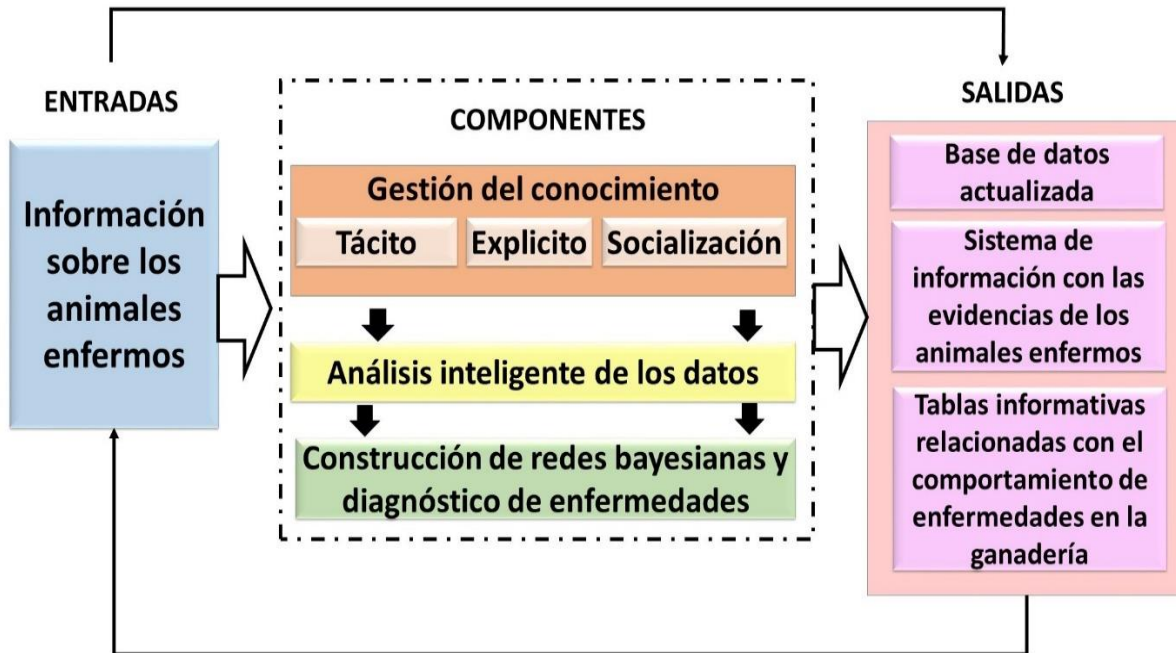
1. La actualización permanente mediante la retroalimentación de la información que nutre al modelo.
2. El enfoque a procesos para definir las actividades necesarias que permitan alcanzar el resultado deseado, identificar las posibles entradas y salidas, así como la evaluación de los diagnósticos de enfermedades.
3. El enfoque hacia la gestión de un conjunto de procesos para identificar, entender y gestionar la información de las enfermedades ganaderas que con frecuencia afectan la ganadería bovina.
4. El enfoque hacia la toma de decisiones, basado en el análisis de la información.

**Los componentes del modelo son:**

1. Gestión del conocimiento
2. Análisis inteligente de los datos
3. Construcción de redes bayesianas y diagnóstico de enfermedades

El modelo está compuesto por componentes relacionados entre sí como se muestra en la figura

2.2.



**Figura 2.2.** Modelo basado en redes bayesianas para el diagnóstico de enfermedades en la ganadería (**MRB-DiagPron**). Fuente: Elaboración propia.

### 2.3.1. Descripción general del modelo

Sobre los principios y cualidades anteriormente establecidos se desarrolló el modelo **MRB-DiagPron** para el diagnóstico de la Fasciolosis bovina utilizando redes bayesianas. Su objetivo es brindar a los especialistas de salud animal información sobre los animales enfermos para diagnosticar la enfermedad que más se ajusta a los síntomas y signos a partir del conocimiento gestionado y analizado. El modelo **MRB-DiagPron** es un modelo útil para apoyar la toma de decisiones por parte de los especialistas de salud animal, con el objetivo de conservar la masa ganadera.

#### **Información de entrada y salida del modelo:**

Las entradas del modelo lo constituyen la información sobre los animales enfermos. Como salida se obtiene base de datos actualizada, sistema de información con las evidencias de los

animales enfermos y tablas informativas relacionadas con el comportamiento de enfermedades en la ganadería.

### **2.3.2. Descripción de los componentes del modelo**

La relación entre los diferentes componentes del modelo se establece a partir de la información que se define como entrada y salida en cada uno de ellos. Cada componente será descrito en función de los algoritmos que se implementan y su formulación matemática.

#### **a) Componente de gestión del conocimiento**

**Entradas:** información sobre las sospechas que se observan en los animales enfermos (síntomas y signos), cantidad de animales susceptibles y enfermos y seguimiento de las sospechas de enfermedades que se observan en los animales enfermos.

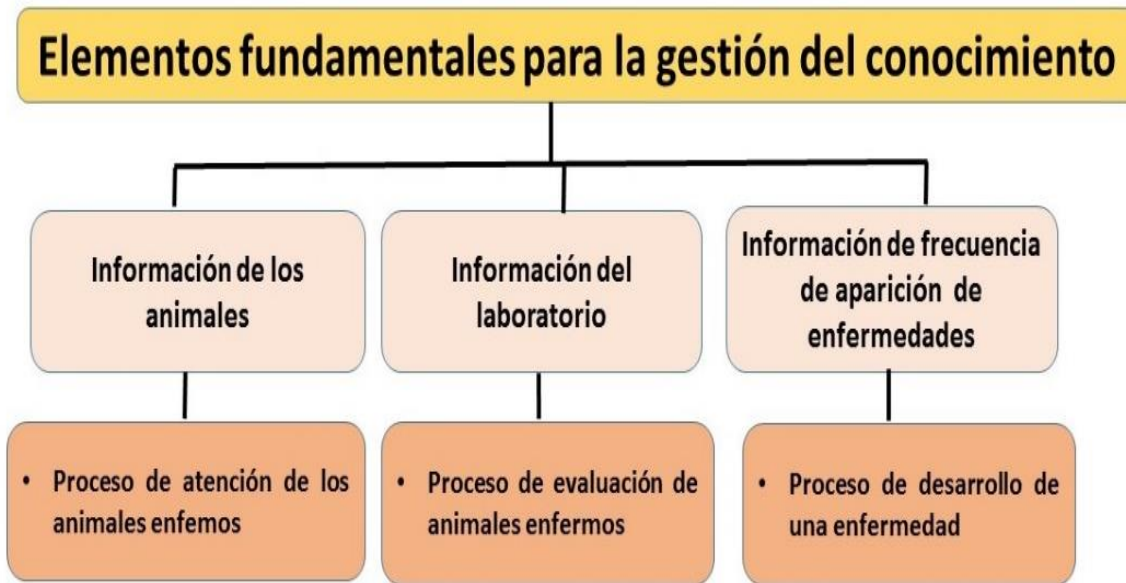
**Salidas:** información enriquecida, asociada a los datos necesarios, para el diagnóstico de enfermedades, en aras apoyar la toma de decisiones.

#### **Descripción general:**

En el modelo **MRB-DiagPron** la gestión del conocimiento constituye un pilar para el desarrollo de los procesos que lo integran pues ellos tienen como base el empleo del conocimiento para la toma de decisiones. Las herramientas utilizadas en el centro objeto de estudio hacen un uso limitado de la información y el conocimiento. Ante esta situación una acción indispensable fue potenciar la gestión de la información y el conocimiento (seleccionarla, organizarla, socializarla a todos los que la necesiten).

La información gestionada es amplia y lo que se trata es de lograr una mayor utilización de la misma para apoyar la toma de decisiones. El componente de gestión del conocimiento describe los elementos fundamentales a tener en cuenta para ser incluidos en los sistemas de información y vigilancia epizootiológica. Esto permite una gestión más adecuada de la información y el

conocimiento, considerado como imprescindible para apoyar la toma de decisiones como se muestra en la figura 2.3.



**Figura 2.3.** Elementos para la gestión del conocimiento. Fuente: Elaboración propia.

En el proceso de atención de los animales enfermos, se gestiona información relacionada con los síntomas que se observan en la masa ganadera enferma. La información del laboratorio se enriquece a partir del proceso de evaluación de los animales enfermos, incrementándose el conocimiento de los especialistas de salud animal para ejecutar el diagnóstico de enfermedades. La información de frecuencia de aparición de enfermedades se gestiona a través de las tasas de incidencia, que mide la probabilidad de que los animales sanos desarrollen enfermedades durante un período de tiempo específico. Representa el número de casos nuevos de enfermedad en una población durante un período de tiempo definido.



$$Incidencia = \frac{\text{número de casos nuevos}}{\text{población en riesgo}} \text{ en un período de tiempo} \quad (2.1)$$

En la tabla 2.1 se muestra la información y conocimiento a gestionar como parte del componente de gestión del conocimiento.

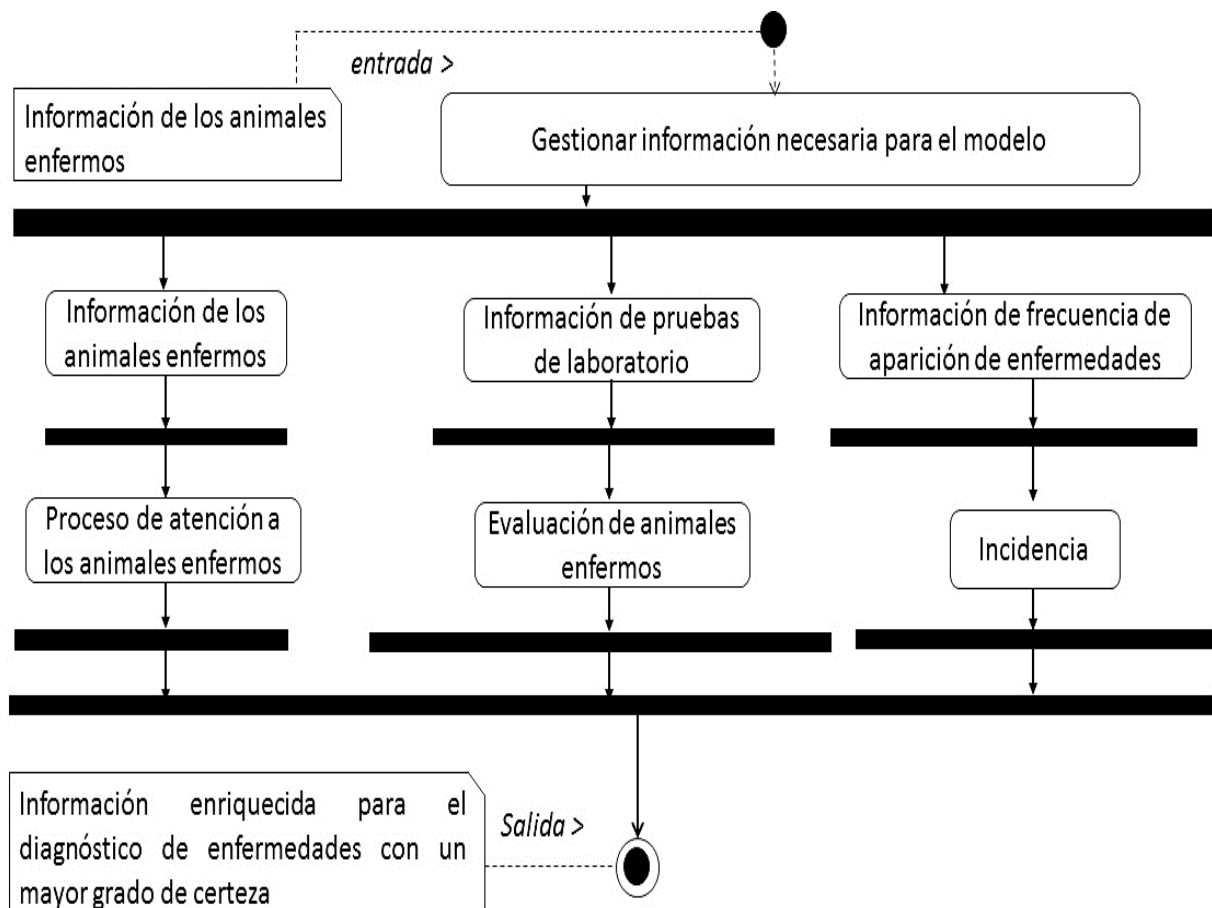
**Tabla 2.1.** Información y conocimiento a gestionar como parte del componente de gestión del conocimiento. Fuente: Elaboración propia.

<b>Categorías del conocimiento</b>	<b>Descripción</b>	<b>Ubicación</b>	<b>Documento utilizado</b>	<b>Información y conocimiento a gestionar</b>
<b>Explicito</b>	Incrementa la eficiencia y eficacia organizacional gestionando el conocimiento, codificándolo a través del uso de las tecnología de la información.	Información de animales enfermos y en la información de laboratorio. Información de frecuencia de enfermedad.	Modelo 622. Focos de enfermedades y modelo de la red diagnóstica, perteneciente al Sistema de Información y Vigilancia Epizootiológica. Modelo de mortalidad perteneciente al Sistema Estadístico Veterinario.	Focos nuevos, recuperados y en situación, las sospechas que se observan en los animales enfermos. Procedimientos, para evolucionar las sospechas que se observan en los animales enfermos.
<b>Tácito</b>	Toma un enfoque de personalización, es aquel que no puede ser fácilmente explicado, es casi propio, intrínseco.	Información de los animales susceptibles a la enfermedad y de posibles muertes por Fasciolosis bovina.		Existencia de casos de nuevos animales enfermos, sobre la población en riesgo, en un período de tiempo. Información que se corresponde con la incidencia de la enfermedad. Número existente de casos sobre la población total en un punto de tiempo.

# Modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina

## Capítulo II

El conocimiento gestionado se socializa a través de la participación activa de los especialistas de salud animal en capacitaciones relacionadas con la gestión de la información para ejecutar diagnóstico de enfermedades, desde los síntomas que se observan en los animales enfermos, en aras de apoyar la toma de decisiones, cuidar la salud animal y conservar la masa ganadera. Proporciona una representación simplificada de la realidad, para obtener mejoras en los diagnósticos de enfermedades en la ganadería. La figura 2.4, representa el proceso de gestión del conocimiento en el modelo **MRB-DiagPron**.



**Figura 2.4.** Proceso de gestión del conocimiento. Fuente: Elaboración propia.

**b) Componente de Análisis inteligente de los datos**

**Entradas:** información de animales enfermos.

**Salida:** registro de datos preprocesados.

**Descripción:**

En el componente de Análisis inteligente de los datos, se selecciona los datos de dos fuentes diferentes:

- Sistema Estadístico Veterinario que tiene una base de datos operacional en Excel.
- Sistema de Información y Vigilancia Epizootiológica, que tiene una base de datos en Access.

Los datos de entrada seleccionados de las fuentes mencionadas se preprocesaron para eliminar los datos que causen contratiempo en los resultados esperados, en aras de garantizar su disponibilidad, completitud y fidelidad.

El preprocesamiento de datos, en la presente investigación, se ejecuta utilizando la tarea de limpieza de datos, para luego incorporar los datos limpios a la base de datos creada previamente. Esta tarea se realiza de forma automática, utilizando el algoritmo **K-Means**.

El algoritmo **K-Means** tienen como base la optimización de una función criterio, donde en la presente investigación, se denomina  $F$ , el valor de esta función depende de las particiones del conjunto de datos  $\{C_1, \dots, C_k\}$

$$F: P_k(X) \rightarrow \mathbb{R} \quad (2.2)$$

Donde:

$P_k(X)$ , son las particiones del conjunto de datos  $X = \{x_1, \dots, x_n\}$  en  $K$  grupos no vacíos.  $x_i$ , es un vector  $n$ -dimensional (objeto) del conjunto de datos  $X$ .

El algoritmo **K-Means** converge a un mínimo local, utilizando la función criterio  $F$ , de la sumatoria de las distancias  $L2$  entre cada objeto y su centroide más cercano. A este criterio normalmente se le denomina error cuadrático y se obtiene a través de la expresión 2.3.

$$F(\{C_1, \dots, C_K\}) = \sum_{i=1}^K \sum_{j=1}^{p_i} \|x_{i_j} - \bar{C}_i\| \quad (2.3)$$

Donde:

$K$  es el número de grupos,  $p_i$  es el número de objetos del grupo  $i$ ,  $x_{i_j}$  es el  $j$ -ésimo objeto del  $i$ -ésimo grupo y  $\bar{C}_i$  es el centroide del  $i$ -ésimo grupo el cual es calculado a través de la expresión 2.4.

$$\bar{C}_i = \frac{1}{p_i} \sum_{j=1}^{p_i} X_{i_j}, i = 1, \dots, K \quad (2.4)$$

El conjunto de pasos lógicos del algoritmo **K-Means** es el que se presenta a continuación:

Paso 1. Selecciona los  $K$  centroides iniciales  $\{C_1, \dots, C_k\}$ .

Paso 2. Asigna los objetos  $x_i$  del conjunto de datos  $X$ , a su centroide más cercano.

Paso 3. Recalcula los nuevos centros, regresa al paso 2, hasta que el algoritmo converge.

El algoritmo se inicia seleccionando o calculando los centroides iniciales, dependiendo del criterio de selección de centroides, posteriormente asigna los objetos a su centroide más cercano, para después recalcular los nuevos centroides esto lo realiza hasta que el algoritmo converja (paso 3).

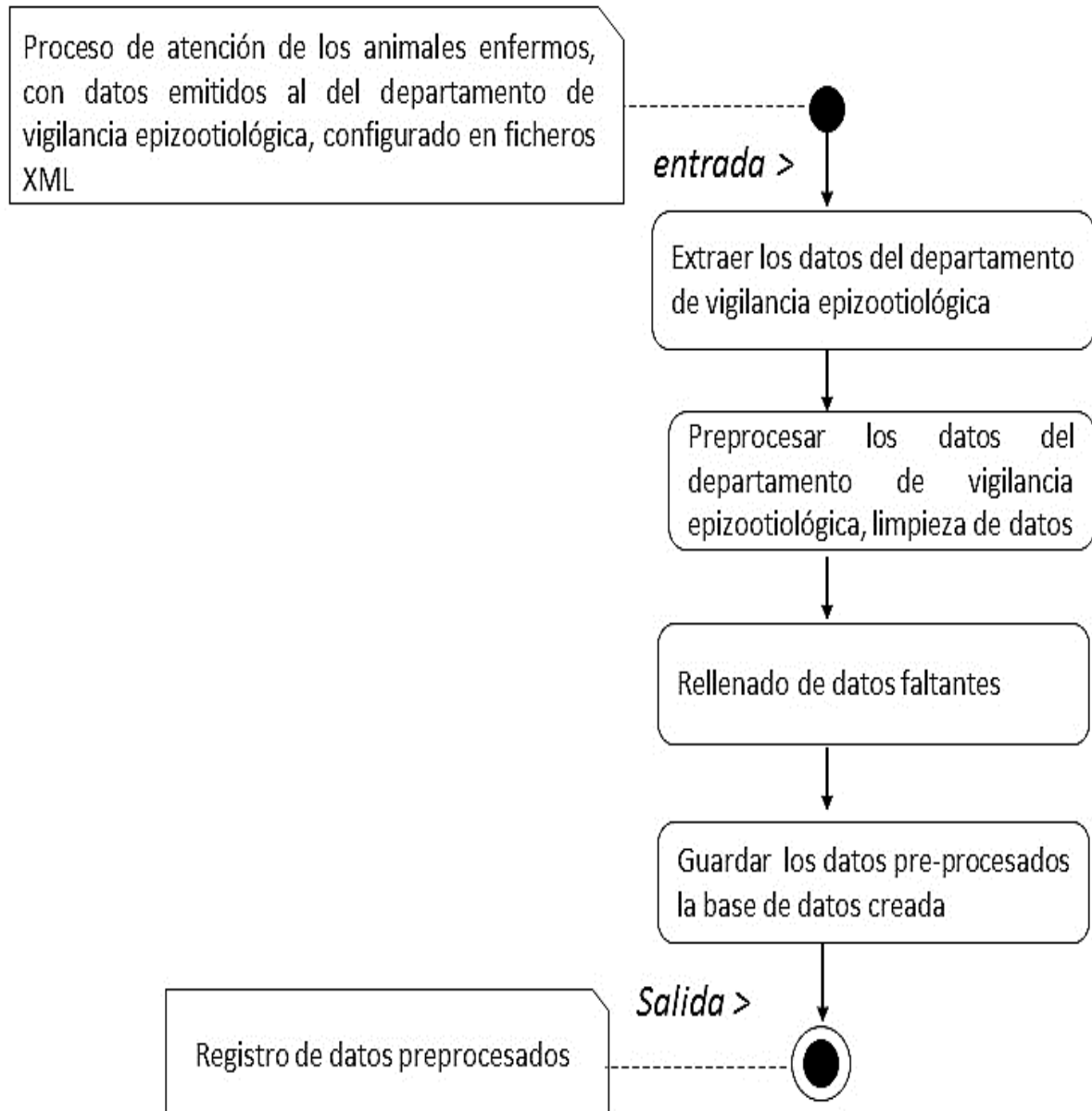
## Modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina

### Capítulo II

Preprocesados los datos y aplicada la técnica de limpieza de datos, los datos faltantes se rellenan, utilizando el método de imputación por media. Método que sustituye los valores faltantes de una variable mediante la media de las unidades observadas en esa variable.

En la presente investigación se realiza una imputación por media condicional, método que imputa medias condicionadas a valores observados y consiste en agrupar los valores observados y no observados en clases e imputar los valores faltantes por la media de los valores observados en la misma clase. El método de imputación de valores faltantes contribuye a reducir la pérdida de los datos faltantes en la base de datos (Castro, 2014).

En la figura 2.5, se representa el proceso de Análisis inteligente de los datos, que se corresponde con las sospechas de los animales enfermos, reportadas a los departamentos de vigilancia epizootiológica de las instituciones agropecuarias.



**Figura 2.5.** Proceso de Análisis inteligente de los datos. Fuente: Elaboración propia.

**c) Componente construcción de redes bayesianas y diagnóstico de enfermedades**

**Entrada:** síntomas y signos de los animales enfermos

**Salida:** diagnóstico de la Fasciolosis bovina

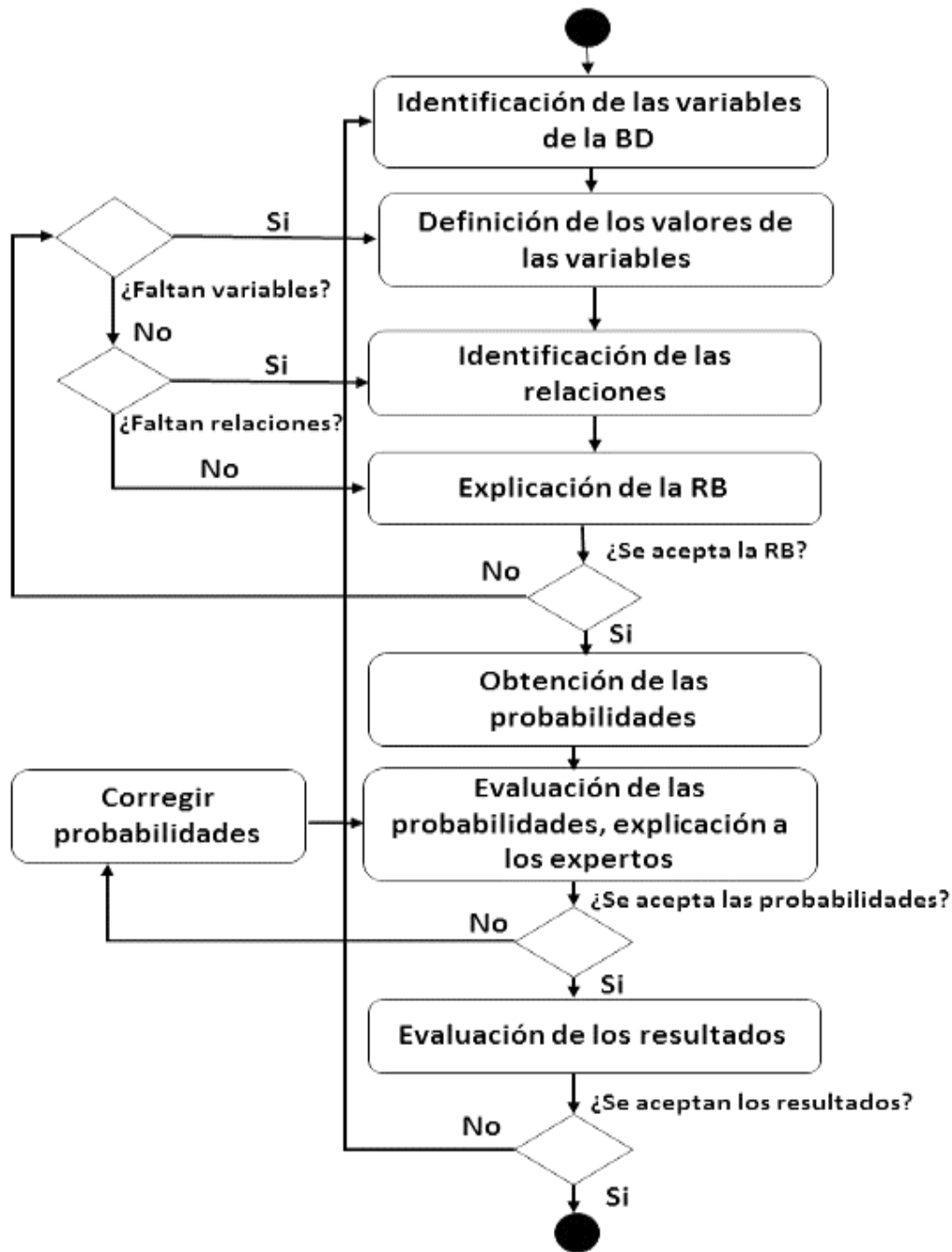
**Descripción:**

Las redes bayesianas son herramientas de modelado estadístico, destinadas a representar un conjunto de incertidumbre relacionada. Su estructura gráfica y su fundamento probabilístico las hace apropiadas para modelar sistemas multivariados orientados al diagnóstico médico para apoyar la toma de decisiones.

El proceso de construcción de redes bayesianas consta de tres fases:

- 1. Definición del grafo;** se identifican las variables del problema y las relaciones de dependencia e independencia entre ellas.
- 2. Identificación de los modelos canónicos;** para ajustar los grupos de variables que se representan en la red bayesiana y reducir el número de parámetros a obtener.
- 3. Obtención de datos cuantitativos;** se obtienen las probabilidades *a priori* de las variables que no tienen padres y las probabilidades condicionadas para el resto de las variables.

El proceso de construcción de redes bayesianas se muestra resumido en la figura 2.6.

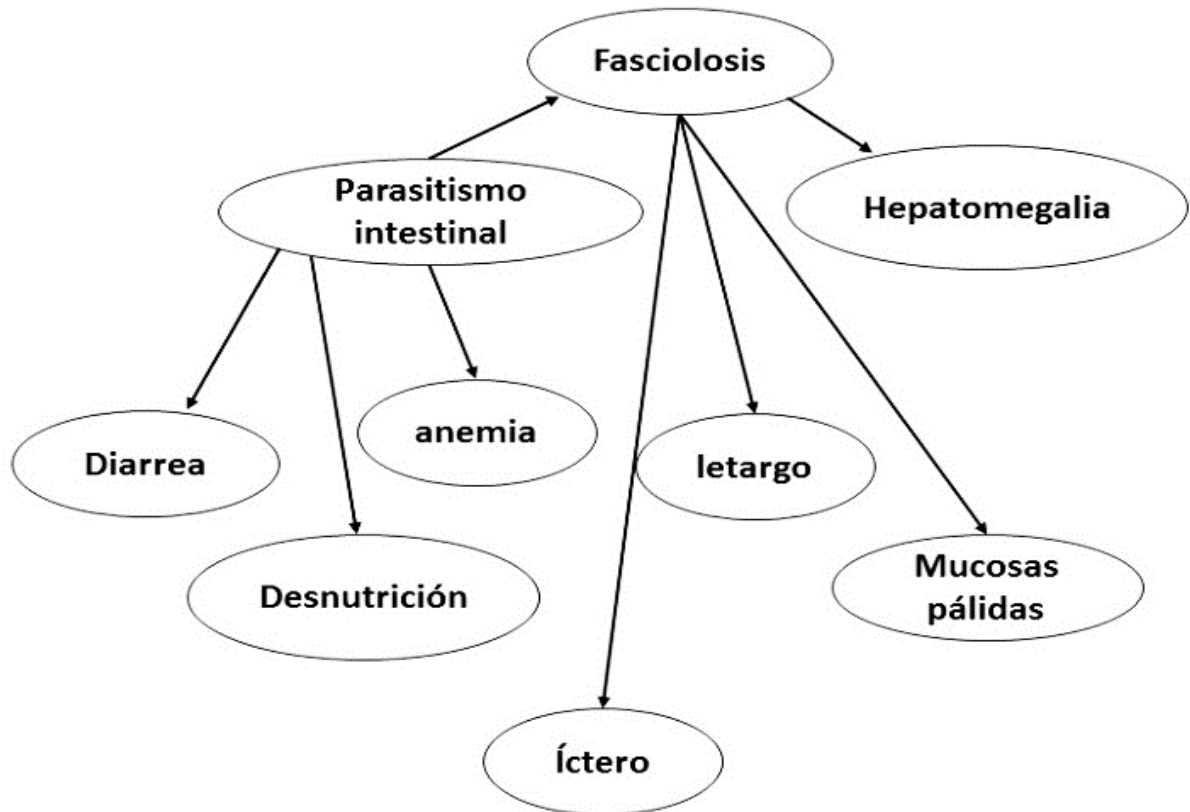


**Figura 2.6.** Algoritmo para la construcción de redes bayesianas. Fuente: Elaboración propia.

En el proceso de construcción de la red bayesiana para el diagnóstico de la Fasciolosis bovina se crean relaciones de inferencia causal entre dos variables. La presencia o ausencia de que una



enfermedad tiene inferencia en que el resultado de las pruebas de laboratorio sea positivo o negativo y definen reglas de diagnóstico del tipo “**si, entonces**”. Los nodos hijos de la red representan los síntomas observados en los animales enfermos y los nodos padres, representan las enfermedades (Fasciolosis o parasitismo intestinal), como se muestra en la figura 2.7.



**Figura 2.7.** Relación padre-hijo, donde *Fasciolosis* resulta ser nodo padre. Fuente:  
Elaboración Propia

La figura 2.7, muestra que la Fasciolosis es la enfermedad parasitaria donde los síntomas con mayor relevancia son: Ícteros (I) y Hepatomegalia (H), por lo que las direcciones en los arcos van dirigidas desde la enfermedad hacia los síntomas, representando una dirección causal. El síntoma de Parasitismo intestinal (PI), es a su vez una enfermedad con síntomas que la

## Modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina

### Capítulo II

identifican y que son comunes para la Fasciolosis, por lo que la dirección del arco es desde el síntoma hacia la enfermedad, representando una dirección de diagnóstico.

Definido el grafo, se introduce la información numérica en la red bayesiana, que se corresponde con la tabla de probabilidad condicionada (TPC) de cada nodo. La TPC se obtiene en base a la experiencia de los médicos y en base a los datos utilizando el teorema de Bayes, donde se debe conocer:

- ✓ **Prevalencia:** cantidad expresada en % de animales que tuvieron la enfermedad en un período determinado.
- ✓ **Sensibilidad:** probabilidad de que el resultado sea positivo, si los animales tienen la enfermedad.
- ✓ **Especificidad:** probabilidad de que el resultado sea negativo, si los animales no tienen la enfermedad.

La información numérica en la red bayesiana para el Parasitismo intestinal, correspondiente a la TPC de ese nodo, se representa en la tabla 2.2.

**Tabla 2.2.** Probabilidad condicionada de Fasciolosis, dado Parasitismo intestinal. Fuente: Elaboración propia.

<b>Parasitismo_intestinal</b>	<b>si</b>	<b>no</b>
<b>Si</b>	0.57	0.47
<b>No</b>	0.43	0.53

Para la variable enfermedad (Fasciolosis), la TPC viene dada con su prevalencia incluida, y el valor probabilístico de Parasitismo intestinal se obtiene de la experiencia de los expertos (especialistas de salud animal). Los cálculos de probabilidad condicional para la Fasciolosis se

## Modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina

### Capítulo II

obtienen a través de la expresión 2.8, para calcular la probabilidad de Fasciolosis dado Parasitismo Intestinal presente. A través de expresión 2.9 se calcula la probabilidad de Fasciolosis dado Parasitismo Intestinal ausente y la expresión 2.10, se utiliza para calcular la probabilidad de Fasciolosis ausente.

$$P(\text{Fasciolosis} = \text{si} | \text{PI}) \quad (2.8)$$

$$P(\text{Fasciolosis} = \text{no} | \text{PI}) \quad (2.9)$$

$$P(\text{Fasciolosis} = \text{no}) \quad (2.10)$$

Donde:

$P(\text{Fasciolosis} = \text{si} | \text{PI})$ , es la probabilidad de que ante la presencia de Parasitismo intestinal la Fasciolosis esté presente.  $P(\text{Fasciolosis} = \text{no} | \text{PI})$ , es la probabilidad de que ausente el Parasitismo intestinal la Fasciolosis esté ausente.  $P(\text{Fasciolosis} = \text{no})$ , es la probabilidad de que esté presente la Fasciolosis.

El cálculo de probabilidades condicionada para la variable síntoma (Íctero), se construye teniendo en cuenta el valor de sensibilidad y la especificidad. La sensibilidad, es la probabilidad de que el síntoma de positivo (si) cuando la enfermedad está presente, y la especificidad, es la probabilidad de que él síntomas de negativo (no) cuando la enfermedad está ausente. La TPC que se obtiene para este síntoma, es la que se muestra en la tabla 2.3.

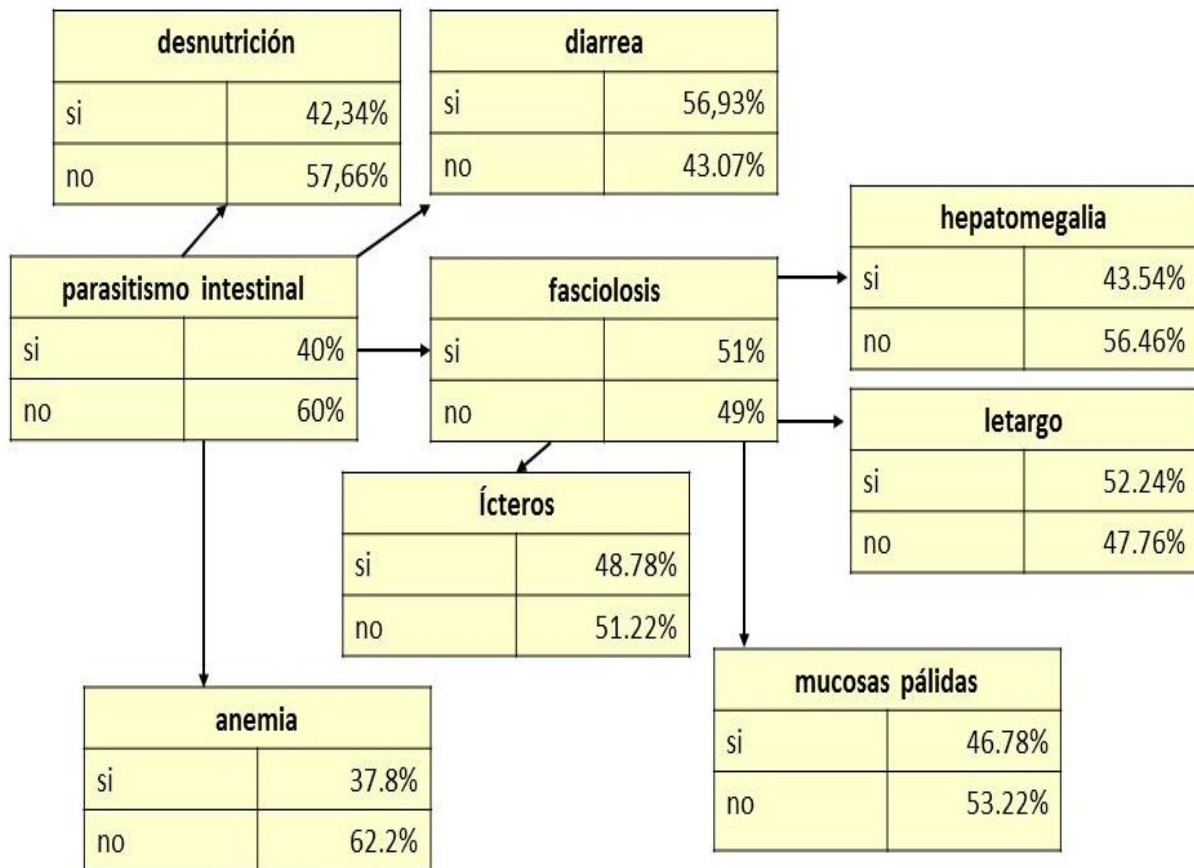
**Tabla 2.3.** Probabilidad condicionada del síntoma Íctero. Fuente Elaboración propia

Íctero	si	no
Si	0.87	0.09
No	0.13	0.91

## Modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina

### Capítulo II

Obtenida las TPC, se realiza la inferencia bayesiana, la que se utiliza para la búsqueda de la hipótesis más factible, dentro de un conjunto de datos y conocimientos previos de la probabilidad de cada hipótesis. Como método bayesiano permite modificar valores cuando se dispone de nueva información. En la figura 2.8, se muestra el resultado de cada nodo al realizar la inferencia bayesiana.



**Figura 2.8.** Inferencia bayesiana. Fuente: Elaboración propia.

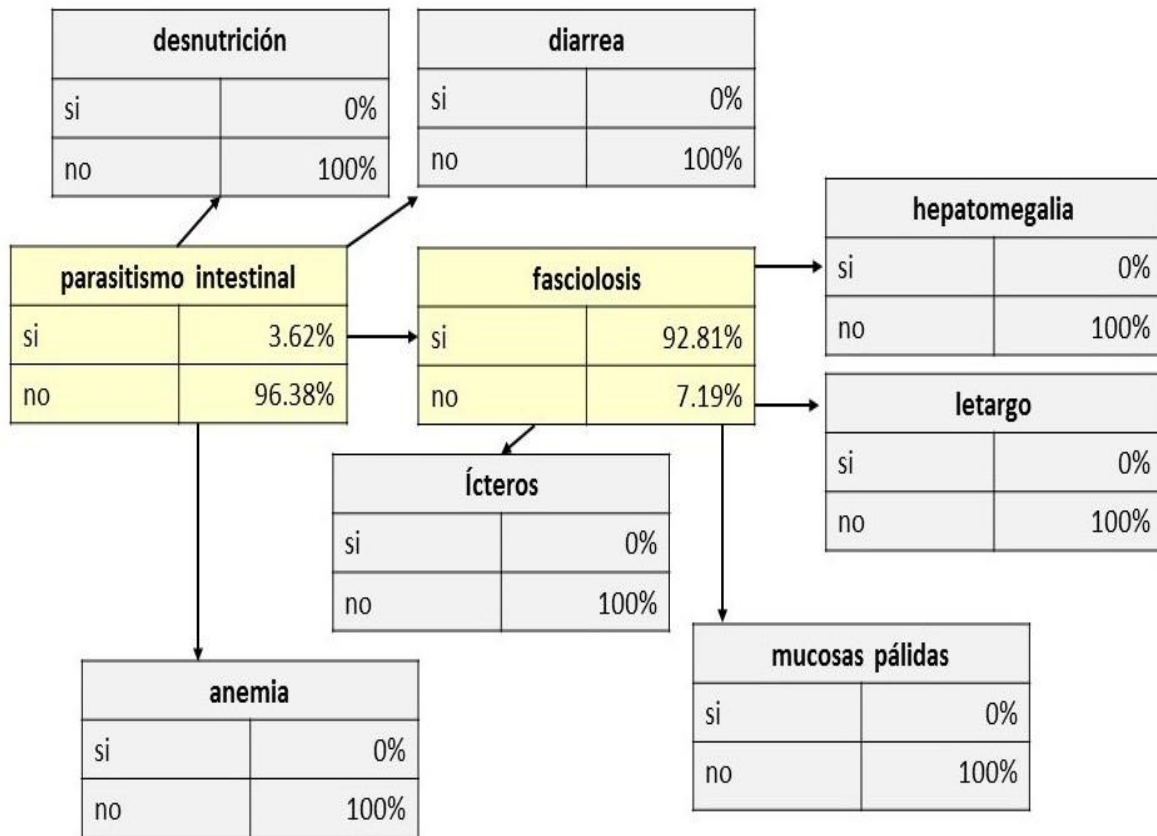
Realizada la inferencia bayesiana, se comprueba que se tiene evidencia de Parasitismo intestinal (PI), útil para el cálculo de la probabilidad a priori de la *fasciolosis* = *si*, cálculo que se realiza a través de la fórmula 2.11 para la parte positiva y 2.12 por la negativa.

$$P(PI = si) = P(PI = si|fasciolosis = si) * P(fasciolosis = si) + P(PI = si|fasciolosis = no) * P(fasciolosis = no) \quad (2.11)$$

$$P(PI = no) = P(PI = no|fasciolosis = si) * P(fasciolosis = si) + P(ícteros = si|fasciolosis = no) * P(fasciolosis = no) \quad (2.12)$$

Para los síntomas, el cálculo de las probabilidades a priori, se realiza igual que para la enfermedad (expresiones 2.8, 2.9 y 2.10). Obtenidos los resultados de las probabilidades a priori para las enfermedades y los síntomas, se introducen hallazgos para obtener las probabilidades a *posteriori* (condicional).

Al contar con las probabilidades a priori y a posteriori, se calcula el valor predictivo positivo (VPP) que representa la certeza con la que se diagnostica la enfermedad, cuando los síntomas están presentes, *si: P(Enfermedad = si | síntoma = si)*, obteniéndose así la enfermedad más probable que presentan los animales enfermos. En la figura 2.9, se representan los resultados de la introducción de hallazgos en la red bayesiana. Los resultados se muestran expresados en porcentos para cada nodo síntoma y para los nodos que representan la enfermedad.



**Figura 2.9.** Introducción de hallazgos en la red bayesiana para Fasciolosis – Parasitismo intestinal. Fuente: Elaboración propia.

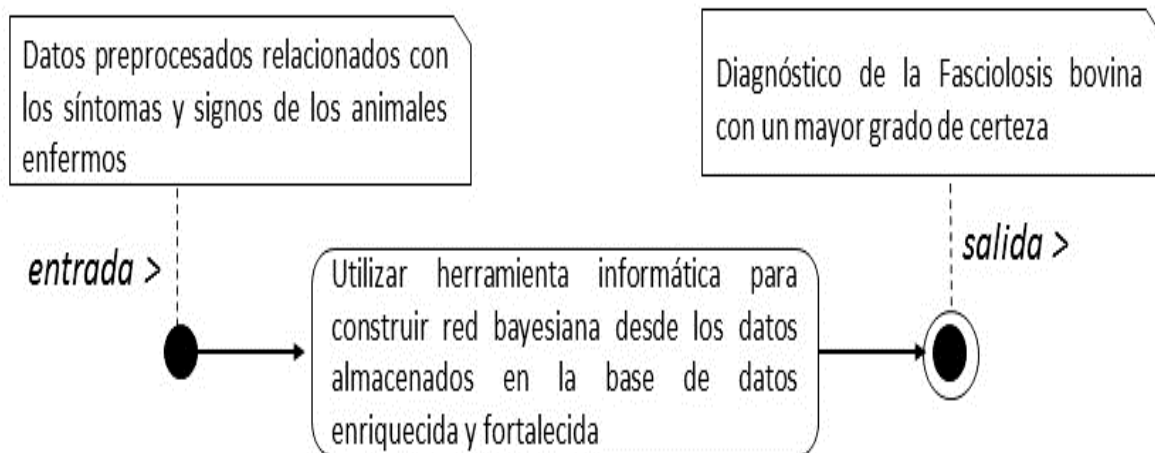
Basado en que existen enfermedades con síntomas comunes, a los de la Fasciolosis, a través del criterio de los expertos, se analiza la enfermedad de Fasciolosis para su diagnóstico, con dos síntomas observables presentes en los animales enfermos. Se calcula el valor predictivo positivos para los síntomas presentes y observables en los animales enfermos (Ícteros y mucosas pálidas) y la enfermedad positiva, expresión 2.13 y el valor predictivo negativo para los síntomas presentes y observables en los animales enfermos (Ícteros y mucosas pálidas) y la enfermedad negativa, expresión 2.14.

$$VPP = P(\text{fasciolosis} = \text{si} | \text{ictero} = \text{si}, \text{mucosas pálidas} = \text{si}) = 92.81\% \quad (2.13)$$

$$VPN = P(\text{fasciolosis} = \text{no} | \text{íctero} = \text{no}, \text{mucosas pálidas} = \text{no}) = 7.19\% \quad (2.14)$$

El valor predictivo positivo (VPP) obtenido para diagnosticar la presencia de Fasciolosis, cuando los síntomas están presentes, resulto ser del 92,81 %, por lo que la enfermedad más probable que presentan los animales enfermos es la Fasciolosis.

La figura 2.10, representa el proceso de construcción de redes bayesianas para el diagnóstico de la Fasciolosis bovina.



**Figura 2.10.** Proceso de construcción de redes bayesianas para el diagnóstico de la Fasciolosis bovina. Fuente: Elaboración propia.

#### 2.4. Conclusiones del capítulo

- El diagnóstico realizado sobre el estado actual de las instituciones agropecuarias, respecto al proceso de gestión y almacenamiento de la información para ejecutar

diagnóstico de enfermedades en la ganadería, permitió constatar las principales dificultades e insuficiencias existentes, así como su importancia e incidencia en la certeza de los diagnósticos de enfermedades ganaderas.

- El proceso de gestión del conocimiento y la socialización del conocimiento tácito y explícito contribuyó a que ese valioso recurso (la información) llegase a todo el que lo necesita y así apoyar la toma de decisiones en el diagnóstico de enfermedades en la ganadería.
- El modelo propuesto para el diagnóstico de la Fasciolosis bovina favorece tiene un impacto positivo en la certeza de los diagnósticos, lo que incide favorablemente en la toma de decisiones.



**Capítulo III: Validación del modelo basado  
en redes bayesianas para el diagnóstico  
de la Fasciolosis bovina**

### **CAPÍTULO III. VALIDACIÓN DEL MODELO PARA EL DIAGNÓSTICO DE LA FASCIOLOSIS BOVINA**

En el capítulo se presenta la herramienta informática como instanciación del modelo **MRB-DiagPron** para el diagnóstico de la Fasciolosis bovina y se abordan las indicaciones metodológicas para su implantación. Se describen los resultados de la validación del modelo a partir del empleo de varios métodos y su aplicación en un entorno real. En el proceso intervinieron expertos, los cuales constataron la hipótesis planteada, donde el modelo mejora la certeza de los diagnósticos de enfermedades en la ganadería. Por último, es analizado el impacto social y factibilidad económica de la propuesta de solución.

#### **3.1. Instanciación del modelo MRB-DiagPron**

La herramienta informática se desarrolló en un entorno Web y se identifica por **DiagPron**. Contribuye a minimizar las dificultades en cuanto a la gestión, procesamiento, almacenamiento y análisis de la información. Esta herramienta se realizó teniendo en cuenta la interacción con los datos almacenados en la base de datos relacionados con las enfermedades ganaderas que mayor afectación ocasionan, como lo es la Fasciolosis bovina. Posee la facilidad extenderse para el análisis de diferentes enfermedades y especies. Brinda la posibilidad de que los médicos veterinarios accedan a ella e introduzcan datos para el análisis sobre el diagnóstico de enfermedades que proporciona (González, Estrada y Romillo, 2016).

**DiagPron** presenta opciones adicionales tales como:

- **Datos proporcionados por los expertos:** se recopila la información que brindan los expertos sobre el comportamiento o la presencia de síntomas de una posible enfermedad, así como las posibles muertes. Esto resulta útil para completar los datos ausentes en la base de datos de acuerdo a lo que observaron a nivel de campo. Con la información que aportan los

expertos y los datos que gestiona el departamento de vigilancia epizootiológica es posible construir redes bayesianas, realizar inferencia y obtener valores predictivos que posibilitan el diagnóstico de enfermedades. Esta funcionalidad está asociada al componente “Gestión del conocimiento” del modelo.

- **Datos estadísticos:** se corresponde con los datos que se generan del Sistema de Información y Vigilancia Epizootiológica y de la estadística que se obtiene sobre focalidad y mortalidad de la masa ganadera los cuales se recopilan y se almacenan en la base de datos que se crea con la herramienta, para utilizarlos en los diagnósticos de enfermedades. Esta funcionalidad está asociada al componente “Análisis inteligente de datos” del modelo.
- **Análisis estadístico:** se corresponde con tablas y gráficos que se obtienen con la herramienta sobre el comportamiento de las enfermedades, brindando gráficos de tendencia, comparación de las medias y cálculos de prevalencia de las enfermedades en diferentes períodos de tiempo. Esta funcionalidad está asociada al componente “Análisis inteligente de datos” del modelo.
- **Construcción de redes bayesianas:** se construyen redes bayesianas para inferir que tan probable sea una causa, útil para realizar el diagnóstico de enfermedades desde los datos almacenados en la base de datos y desde los datos que aportan los expertos. Esta funcionalidad está asociada al componente “Construcción de redes bayesianas y diagnóstico de enfermedades” del modelo.
- **Explicación de la red:** proporciona una explicación detallada de la red bayesiana construida para el diagnóstico de enfermedades. Esta funcionalidad está asociada al componente “Construcción de redes bayesianas y diagnóstico de enfermedades” del modelo.

### **3.2. Funcionamiento general de la aplicación**

La herramienta desarrollada para el diagnóstico de enfermedades en la ganadería (**DiagPron**), tiene la ventaja de interactuar con las bases de datos, relacionadas con la vigilancia epizootiológica, las cuales contienen información sobre el seguimiento de las enfermedades ganaderas, que mayor afectación ocasionan. Los datos se gestionan, se preprocesan, se analizan y se almacenan en la base de datos de **DiagPron**, facilitando la construcción de redes bayesianas y la ejecución de los diagnósticos de enfermedades.

La información que se trata, desde las bases de datos, relacionadas con la vigilancia epizootiológica es:

- Notificaciones de las sospechas de animales enfermos, incluye: (Fecha, Provincia, Municipio, Especie y Síntomas).
- Cantidad de animales susceptibles.
- Cantidad de animales enfermos.
- Cantidad de animales muertos.
- Datos relacionados con la evolución de las sospechas de los animales enfermos, incluye nombre de la enfermedad.

Los datos antes referidos se gestionan en el componente “Gestión de conocimiento” y se preprocesan en el componente “Análisis inteligente de datos”, a través de la herramienta informática creada para el diagnóstico de enfermedades en la ganadería. Luego se almacena en la base de datos creada con **DiagPron**, obteniéndose una base de datos con información limpia. El marco de trabajo para construir la interfaz de la herramienta se seleccionó a partir de la interacción de los componentes de gestión del conocimiento, Análisis Inteligente de los Datos, Construcción de redes bayesianas y diagnóstico de enfermedades en la ganadería.

**DiagPron**, está diseñada bajo los principios de interoperabilidad, permitiendo que la información que viaje desde el punto donde se proporciona hasta el punto donde se consume se estandarice, sin importar las diferencias de la tecnología en que están desarrolladas las aplicaciones consumidoras. Se ha comenzado a utilizar como herramienta de apoyo a la toma de decisiones para el cuidado y conservación de la masa ganadera (González, Estrada y Romillo, 2016).



**Figura 3.1.** Vista principal de la aplicación. Fuente: Elaboración propia.

### **3.3. Indicaciones metodológicas para la aplicación del modelo**

Las indicaciones metodológicas para el diagnóstico de enfermedades en la ganadería, ofrecen una guía para la implantación del modelo. En la Figura 3.1 se muestran las etapas propuestas en las indicaciones metodológicas.



**Figura 3.1.** Etapas de la implementación del modelo. Fuente: Elaboración propia.

Las acciones asociadas a cada etapa propuesta son:

- **Diagnóstico:** parte de la aceptación por las instituciones agropecuarias de las premisas para la aplicación del modelo. Se identifica la situación existente para la aplicación del modelo y el despliegue de la herramienta informática.
- **Implantación:** se despliega la herramienta informática y se gestiona toda la información relacionada con la vigilancia epizootiológica de la masa ganadera.
- **Configuración:** se ajusta la configuración de la herramienta de acuerdo a la situación existente.
- **Capacitación:** se enseña al personal decisor en la utilización de la herramienta informática y funcionamiento del modelo. Es una etapa fundamental para apropiar todos los conocimientos necesarios para la implementación del modelo.
- **Acompañamiento:** por un período de tiempo determinado se brinda asesoría al personal de salud animal en el uso de la herramienta informática y entendimiento del modelo.
- **Soporte:** es la etapa final de la implantación, en esta etapa, es donde se realizan los ajustes necesarios al modelo, con el objetivo de realizar diagnóstico de enfermedades en la ganadería con un mayor grado de certeza.

### **3.4. Diseño de la validación**

La validación del modelo se realizó empleando métodos cuantitativos, cualitativos y experimentales. De esta forma se validaron los aspectos relacionados con la certeza de los diagnósticos y la incidencia del modelo en estos resultados, así como los componentes del modelo. Para la validación se emplearon las siguientes técnicas y métodos:

- **Criterio de expertos empleando la escala de *Osgood***: utilizado para validar el modelo y sus componentes.
- **Experimento**: utilizado para realizar la comparación del diagnóstico de la Fasciolosis bovina, antes y después de aplicar el modelo, se evaluó la capacidad del modelo para mejorar la certeza en los diagnósticos de enfermedades en la ganadería. Se aplicó la herramienta informática como instanciación del modelo.
- **Entrevista en profundidad**: utilizada para valorar los beneficios y aportes del modelo, **MRB-DiagPron**, así como conocer opiniones de los especialistas de salud animal con experiencia en la gestión de la información para ejecutar el proceso de diagnóstico de enfermedades en la ganadería, conocimiento empleado para mejorar el mismo.

Finalmente, se empleó la triangulación metodológica inter-métodos para minimizar el sesgo de la investigación.

### **3.5. Valoración de los expertos sobre el modelo**

Se utilizó la escala de *Osgood* para obtener una valoración de 20 especialistas con más de 10 años de experiencia en el trabajo con la salud animal. Los especialistas seleccionados aportaron una valoración sobre los resultados de la investigación realizada.

Las escalas son instrumentos de recolección de información con base en una lista de ítems, reactivos, o frases cuidadosamente seleccionados, de forma que constituyen un criterio

sistemático, confiable, válido y específico para medir cuantitativamente alguna forma de fenómeno. Se evalúa, en este caso, las opiniones de los especialistas de salud animal sobre la validez de los resultados del modelo propuesto en la presente investigación.

En este caso se aplica para validar el resultado de una investigación ya concluida y que ha sido observada directamente por los evaluadores, con la escala que se explica a continuación:

**Muy satisfactorio** 5 4 3 2 1 **Insatisfactorio**

A los especialistas de salud animal se le pidió que evaluaran de la siguiente forma:

- 5 Muy satisfactorio
- 4 Satisfactorio
- 3 Neutro
- 2 Poco satisfactorio
- 1 Insatisfactorio

En la tabla 3.1, se muestran las opiniones de 20 especialistas de salud animal respecto a 6 aspectos relacionados con la investigación. Para ello se tuvo en cuenta los años de experiencia laboral y experiencia en el área de la salud animal, así como labor que desempeña en las instituciones agropecuarias.



**Tabla 3.1.** Resumen de las valoraciones de escala realizada por especialistas de salud animal sobre los resultados de la investigación realizada. Fuente: Elaboración propia.

<b>Ítem a evaluar</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
La manera en que se tratan los diagnósticos de enfermedades en la ganadería en general y en particular los diagnósticos de enfermedades, basado en redes bayesiana en el modelo propuesto.				6	14
Lo beneficioso de utilizar las TIC en el desarrollo de los diagnósticos de enfermedades en la ganadería.			2	6	12
Importancia de las acciones planificadas en el modelo para el diagnóstico de enfermedades en la ganadería.			1	5	14
La concepción general del modelo para el diagnóstico de enfermedades en la ganadería.				5	15
<b>Impacto del modelo en el diagnóstico de la enfermedad de Fasciolosis bovina, con un mayor grado de certeza.</b>				2	18
Opinión sobre las indicaciones metodológicas para implementar el modelo en las instituciones agropecuarias.				6	14
Resultados visibles con la aplicación del modelo en los departamentos de Sanidad Animal.				3	17
<b>Suma total de la valoración</b>			<b>3</b>	<b>33</b>	<b>104</b>

Los datos representados en la tabla 3.1, revelan una buena evaluación de los resultados de la investigación, lo cual se representa en la figura 3.2.

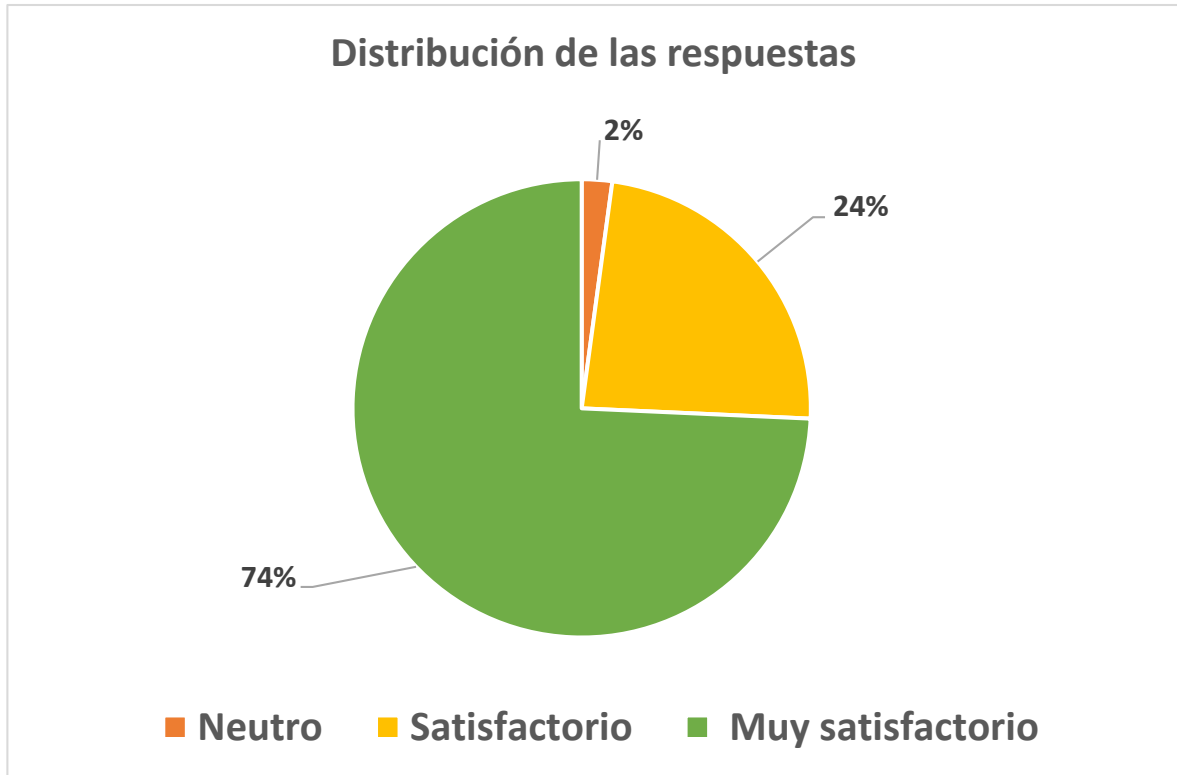


Figura 3.2. Respuestas en la escala de *Osgood*. Fuente: Elaboración propia.

Para calcular el diferencial semántico se escogieron otros pares de adjetivos opuestos y 5 posibilidades de seleccionar, para evaluar la aceptación del modelo diseñado.

**Ítem a evaluar:** El modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis bovina con un mayor grado de certeza

**Espacio semántico:** 1, 2, 3, 4, 5

**Buena ----- Mala** ➡ Los 20 especialistas en salud animal marcaron la primera opción.

**Adecuada ----- Inadecuada** ➡ Los 20 especialistas en salud animal marcaron la primera opción.

<b>Aplicable</b> ----- <b>Inaplicable</b>	➡	Los 20 especialistas en salud animal marcaron la primera opción.
<b>Innovadora</b> ----- <b>Obsoleta</b>	➡	Los 20 especialistas en salud animal marcaron la primera opción.
<b>Completa</b> ----- <b>Incompleta</b>	➡	19 la primera opción, un especialista en salud animal la segunda.

Estos valores son una evidencia de la evaluación favorable del modelo basado en redes bayesianas para el diagnóstico de enfermedades en la ganadería, en particular para la Fasciolosis bovina.

### **3.6. Desarrollo del experimento para el estudio del comportamiento de la Fasciolosis bovina**

Se realizó un experimento para el estudio acerca del comportamiento de la Fasciolosis bovina, que permitió comprobar el cumplimiento de la hipótesis científica de la investigación. Se midió la cantidad de animales enfermos y muertos, antes y después de aplicar el modelo, la certeza del diagnóstico de la Fasciolosis bovina y el comportamiento del decomiso de hígados, en el período de 2012 -2016. Se estratifica el período en dos etapas, una de 2012 – 2014, período antes de aplicar el modelo y la otra de 2015 – 2016, período después de aplicado el modelo.

- **Análisis estadístico de los datos arrojados en el experimento:**

Para evaluar la posibilidad de comparar los datos entre el grupo de animales enfermos y muertos antes de aplicar el modelo y después de aplicarlo para los **indicadores**: decomiso de hígados por Fasciolosis bovina y el diagnóstico de la Fasciolosis bovina, se aplicó la prueba de normalidad de *Shapiro-Wilk* para comprobar que los datos se ajustan a una distribución Normal. Los valores de probabilidad mayores que 0.05 en todos los casos, indican que los datos de las variables se ajustan adecuadamente a la distribución Normal, permitiendo entonces la

aplicación de la prueba *t de Student* para comparar las muestras de cada pareja de variables. En la tabla 3.2 se presenta los resultados del análisis de normalidad de los datos.

**Tabla 3.2.** Análisis de normalidad de las variables según prueba de Shapiro-Wilk. Fuente: Elaboración propia.

Indicadores	Período de las pruebas	Prueba de normalidad para los datos de animales enfermos		Prueba de normalidad para los datos de animales muertos	
		Estadístico W	p-valor	Estadístico W	p-valor
Decomiso de hígados por Fasciolosis bovina	Antes de aplicar el modelo	0,90	0,3952	0,92	0,5162
	Después de aplicar el modelo	0,92	0,5019	0,91	0,4344
Diagnóstico de la Fasciolosis bovina	Antes de aplicar el modelo	0,87	0,2805	0,96	0,7995
	Después de aplicar el modelo	0,93	0,6093	0,95	0,7242

En consecuencia, se aplicó la prueba estadística paramétrica *t de Student*, con el objetivo de comparar dos muestras a partir de sus medias para comprobar si los resultados son estadísticamente diferentes o no. En la tabla 3.3, se muestran los resultados obtenidos para la comparación de las dos muestras según esa prueba estadística.

**Tabla 3.3.** Resultados de la comparación de las muestras según la prueba *t de Student*. Fuente: Elaboración propia.

Indicadores	Animales enfermos						Animales muertos					
	Antes del modelo		Después del modelo		Prueba t		Antes del modelo		Después del modelo		Prueba t	
	Media	Desv. Est.	Media	Desv. Est.	T	P-valor	Media	Desv. Est.	Media	Desv. Est.	T	P-valor
Decomiso	72998,40	2013,44	14599,60	2044,57	45,51	0,0000	773,80	28,60	146,80	9,36	46,60	0,0000
Diagnóstico	53270,20	2114,69	9854,00	3569,71	23,40	0,0000	706,80	22,40	141,40	11,33	50,37	0,0000

Para cada indicador se observa que las medias de animales enfermos y muertos antes y después de aplicado el modelo difieren estadísticamente, dado por los valores de probabilidad menores

que 0,05 en la prueba *t de Student*, lo que significa que con la aplicación del modelo se reducen las cantidades de animales enfermos y muertos, tanto en el decomiso de hígado como para el diagnóstico de la enfermedad.

### **3.7. Entrevista en profundidad**

La entrevista en profundidad supone una conversación con fines orientados a los objetivos de una investigación social (Marcos et al., 2014; Valles, 2014).

La entrevista tiene las siguientes características:

- Discurso, cuyo orden puede resultar más o menos determinado, según sea la reactividad del entrevistado y el flujo de un tema a otro.
- Puntos de referencia de paso obligatorio para el entrevistador y entrevistado.
- Información controlada.
- Información recogida en un mayor tiempo.
- Nivel medio de información previa.

Para el desarrollo de la entrevista el entrevistador utilizó un guión con elementos derivados del problema general a estudiar. A lo largo de la entrevista el entrevistado fue proporcionando información en relación a los elementos necesarios.

Se le realizó entrevistas en profundidad a un conjunto de especialistas en salud animal con más de 20 años de experiencia laborando en los servicios de veterinaria, con altas responsabilidades en los aspectos epizootiológicos. Participaron en las entrevistas en profundidad un conjunto de técnicos veterinarios que trabajan a nivel de campo directamente con la masa ganadera, con más de 10 años en la actividad.

Todos acompañaron el desarrollo y aplicación del modelo **MRB-DiagPron** por un período de un año. Se revisó con detenimiento su principio de funcionamiento una vez terminado, realizando grandes aportes a la investigación.

El Anexo 9 muestra los temas abordados en las entrevistas, entre los principales resultados obtenidos destacan:

- La novedad de la investigación.
- Alta correspondencia de la investigación hacia las tendencias internacionales en los diagnósticos de enfermedades en la ganadería.
- La validez del modelo MRB-DiagPron para el diagnóstico de enfermedades en la ganadería.
- Los beneficios sociales y económicos que aporta la aplicación en el Sistema Nacional de Veterinaria.
- La adecuación de los principios que sustentan el modelo para un fácil y entendible uso por parte de personal no experto en las nuevas tecnologías.

### **3.8. Resultados de la triangulación metodológica de los métodos aplicados**

La triangulación metodológica es una técnica usada para tomar múltiples puntos de referencia y localizar una posición desconocida. Disminuye el sesgo que se produce al comparar resultados obtenidos en la cuantificación de variables mediante un método cuantitativo, las tendencias y dimensiones que surgen de la aplicación de métodos cualitativos (Valencia, 2013). La triangulación metodológica es definida también como la combinación de múltiples métodos en un estudio del mismo objeto o evento para abordar mejor el fenómeno que se investiga (Hussein, 2015).

A partir de la aplicación de los métodos cuantitativos y cualitativos anteriormente expuestos, se realiza la triangulación metodológica inter-métodos de los resultados. La misma permite contrastar los resultados obtenidos, de manera que se puedan determinar las coincidencias y divergencias. Constituye un criterio integrador sobre la validez de la propuesta de solución presentada. El resultado de su aplicación se muestra en la tabla 3.4.

**Tabla 3.4.** Resultados de la triangulación metodológica inter-métodos. Fuente: Elaboración propia.

<b>Objetivo a evaluar</b>	<b>Métodos cuantitativos</b>	<b>Métodos cualitativos</b>	<b>Conclusión</b>
Desarrollo de un modelo, basado en redes bayesianas, para el diagnóstico de la Fasciolosis bovina, capaz de brindar información con un mayor grado de certeza para apoyar la toma de decisiones.	<ul style="list-style-type: none"> <li>- Escala de <i>Osgood</i> (74 % muy satisfactorio)</li> <li>- Experimento</li> </ul>	<ul style="list-style-type: none"> <li>- Entrevista a profundidad</li> </ul>	Validez de la propuesta del modelo para el diagnóstico de la Fasciolosis bovina, capaz de brindar información con un mayor grado de certeza para apoyar la toma de decisiones.

### **3.9. Comparación entre los modelos desarrollados para el análisis del comportamiento de la Fasciolosis bovina**

Se realizó una comparación entre los modelos desarrollados por la autora. Los resultados que se obtienen están relacionados con el comportamiento de la Fasciolosis bovina, desde los datos que reportan al departamento de vigilancia epizootiológica.

#### **Datos empleados para la comparación entre los modelos:**

Se emplearon los modelos desarrollados por la autora para la enfermedad de la Fasciolosis bovina. El modelo de diagnóstico de la Fasciolosis bovina, basado en redes bayesianas y el

modelado del comportamiento de la Fasciolosis bovina, (se analiza la cantidad de animales enfermos y muertos antes y después de aplicar el modelo utilizando redes bayesianas). El modelo del comportamiento de la Fasciolosis bovina, en cuanto a las pérdidas de hígados que esta enfermedad ocasiona y que se reporta por el Sistema Estadístico Veterinario (se analiza el decomiso de hígados, de los animales enfermos y muertos, una vez que son llevados a matadero, antes y después de aplicar diagnóstico con redes bayesianas).

**Criterios de evaluación:**

En el modelo de diagnóstico de la Fasciolosis bovina, basado en redes bayesianas, se evalúa cuando se realiza la inferencia bayesiana y se obtienen resultados, los cuales facilitan evaluar el grado de certeza, en cuanto a la mejora de los diagnósticos, midiéndose a través de los indicadores de cantidad de animales enfermos y muertos antes de aplicar el modelo y después de aplicar el modelo. En el modelo del comportamiento de la Fasciolosis bovina, en cuanto a las pérdidas de hígados que esta enfermedad ocasiona, se evalúa la prevalencia de la enfermedad, cuyo cálculo se realiza teniendo en cuenta la cantidad de hígados decomisados que se corresponden con aquella población de animales sacrificados por la enfermedad de Fasciolosis bovina y la cantidad de bovinos totales, que son los que se corresponden con la masa total de animales enfermos, en el período de análisis.

**Análisis de los resultados**

En la tabla 3.5, se muestran los resultados de la comparación relativo a los aspectos que se tuvieron en cuenta al trabajar con ambos modelos. La comparación realizada antes y después de aplicar el modelo, demostró, que los mejores resultados se obtienen, después de aplicado el modelo, para las medidas de cantidad de animales enfermos y muertos, con un grado de certeza de aproximadamente el 20 %.



**Tabla 3.5.** Resultados de la comparación entre los modelos. Fuente: Elaboración propia.

Modelo	Datos		Variación variable	Población total	Medida	
	Número de conceptos	Número de Experto			Cantidad de animales enfermos	Cantidad de animales muertos
Comportamiento de la Fasciolosis bovina, en cuanto a las pérdidas de hígados que esta enfermedad ocasiona (Prevalencia de la enfermedad)	3	5	antes	1110030	364992 (32% con respecto a la población total)	3669 (0,33% con respecto a la población total)
			después		72998 (6,5 % con respecto a la población total). Disminuyó la cantidad de animales enfermos después de aplicar el modelo en un 19.99%	734 (0.066 % con respecto a la población total). Disminuyó la cantidad de animales muertos después de aplicar el modelo en un 20 %
Diagnóstico de la Fasciolosis bovina, basado en redes bayesianas (Diminución de animales enfermos y muertos)	3	5	antes	1110030	246351 (22.19% con respecto a la población total)	3534 (0,31% con respecto a la población total)
			después		49270 (4,44 % con respecto a la población total). Disminuyó la cantidad de animales enfermos después de aplicar el modelo en un 19,7%	707 (0.066 % con respecto a la población total) Disminuyó la cantidad de animales muertos después de aplicar el modelo en un 20 %

**3.10. Comparación de la red bayesiana para el diagnóstico de la Fasciolosis bovina con las redes bayesianas existentes para uso veterinario**

Durante la caracterización del estado actual sobre el diagnóstico de enfermedades en la ganadería y en particular para la Fasciolosis bovina, en lo relativo a la gestión del conocimiento, análisis inteligente de los datos y la certeza de los diagnósticos de enfermedades, se identificaron cinco redes bayesianas desarrolladas para uso veterinario. Las más importantes son; las desarrolladas para el análisis de datos complejos de enfermedades en los animales, en particular la red bayesiana diseñada para el diagnóstico de la diarrea bovina, la red bayesiana para el diagnóstico de enfermedades tropicales bovinas, la red bayesiana para la especie porcina, específicamente para evaluar los riesgos de infección por micoplasma en cerdos, la red bayesiana para el análisis de la fiebre porcina en manadas y la red bayesiana para el diagnóstico de enfermedades de las especies caninas y felinas.

Las redes bayesianas identificadas en la bibliografía, son poco reutilizables, debido al diseño para el cual fueron concebidas, ellas fueron desarrolladas para especies animales, específicas, donde no es posible la interacción con los expertos, su construcción se realiza de forma manual. Por tal motivo, surge la necesidad de desarrollar una herramienta informática de soporte al modelo **MRB-DiagPron**.

En la tabla 3.6, se presenta la comparación de las redes bayesianas para el diagnóstico de la Fasciolosis bovina, con las redes bayesianas construidas para uso veterinario, reportadas en la literatura, teniendo en cuenta la facilidad de uso y la certeza, las cuales constituyen cualidades del modelo basado en redes bayesianas para el diagnóstico de enfermedades en la ganadería y en particular para el diagnóstico de la Fasciolosis bovina.

Las principales redes bayesianas diseñadas para uso veterinario, reportadas en la bibliografía, presentan dos limitantes fundamentales, en lo relativo a la sensibilidad con respecto a la incoherencia de los datos, lo que provoca poca certeza en los resultados que se obtienen. En cuanto a facilidad de uso están limitadas para el dominio que fueron diseñadas por lo que ni pueden aplicarse para el estudio de otras enfermedades ganaderas.

**Tabla 3.6.** Comparación de las redes bayesianas para uso veterinario. Fuente: Elaboración propia.

<b>Redes bayesianas para uso veterinario</b>	<b>Facilidad de uso</b>	<b>Certeza</b>
Red bayesiana como herramienta para el análisis de datos complejos de animales; diseñada para conocer el comportamiento de la diarrea bovina, (Lewis <i>et al.</i> , 2011).	Facilidad de uso, limitado en los diagnósticos, está diseñada para el análisis de datos complejos de animales, a través de ella sólo se puede conocer el comportamiento de la diarrea bovina.	Sensible con respecto a la incoherencia de los datos, lo que provoca poca certeza en los valores que se obtienen con respecto al comportamiento de la enfermedad para la que fue diseñada.

<p>Redes bayesianas para el diagnóstico de enfermedades tropicales bovinas, (McKendrick <i>et al.</i>, 2000).</p>	<p>Facilidad de uso, limitada ya que es utilizada sólo para el comportamiento y la aparición de enfermedades tropicales bovinas.</p>	<p>Sensible al diagnóstico de enfermedades tropicales de los bovinos. Solo permite utilizar redes bayesianas simples para representar el comportamiento de las enfermedades tropicales lo que provoca que el grado de certeza en los resultados que brinda no sean los adecuados para reutilizarla en otros análisis.</p>
<p>Redes bayesianas para evaluar los riesgos de una infección por micoplasmas en la especie porcina, (Otto y Kristensen, 2004).</p>	<p>Facilidad de uso, limitada se utiliza para la especie en cuestión y para evaluar riesgos de dicha enfermedad.</p>	<p>Sensible a los valores de riesgo de una enfermedad debido a la inconsistencia de los datos, lo que provoca poca certeza en los resultados que ofrece.</p>
<p>Redes bayesianas para el análisis de fiebre porcina en las pjaras, (Geneen y Van der Gaag, 2005).</p>	<p>Facilidad de uso, limitada sólo se utiliza para el análisis de la fiebre porcina.</p>	<p>Sensible a los valores para el análisis de la fiebre porcina, debido a la versatilidad en los datos, lo que causa poca certeza en los resultados que ofrece.</p>

<p>Red bayesiana para el diagnóstico de la Fasciolosis bovina, (González, 2017).</p>	<p>Se obtienen redes bayesianas útiles para el análisis de los resultados que brinda, en particular para el diagnóstico de enfermedades en la ganadería facilitan los análisis, es fácil de utilizar y puede reutilizarse para el estudio de todas las enfermedades ganaderas.</p>	<p>Aumenta la fiabilidad del modelo al emitir valores ciertos para el diagnóstico de enfermedades en la ganadería.</p>
--	--	--

Como se puede apreciar en la comparación, de la red bayesiana para el diagnóstico de la Fasciolosis bovina y las redes bayesianas desarrolladas para uso veterinario, se superan los resultados obtenidos, en cuanto a la combinación de facilidad de uso y certeza, en la red bayesiana desarrollada para el diagnóstico de la Fasciolosis bovina.

**Análisis estático:**

Las propuestas existentes se basan mayormente en el análisis a partir de los datos, o la prescripción de los síntomas, signos y factores de riesgo de las enfermedades. Este hecho limita la flexibilidad del análisis estático a determinadas circunstancias. La certeza también es afectada debido a que la determinación de los nodos más importantes del modelo es un problema multicriterio, (Jun *et al.*, 2010), por lo que lo limitado de los criterios puede llevar a una decisión errónea y ofrecer poco grado de certeza en los resultados que se obtienen.

**3.11. Impacto del modelo**

✓ La cantidad de animales enfermos y muertos por Fasciolosis bovina, disminuyó en el

período 2012 – 2017. El número de focos registrados en el período fue de un 20 % menor que en el período de 2006 – 2011, cuya reducción tiene un impacto positivo en la salud animal de los bovinos.

- ✓ Los especialistas de salud animal a partir de los resultados obtenidos con el modelo **MRB-DiagPron**, han mejorado la atención de salud de los bovinos de la zona (Llanura Sur – Oriental) de la provincia de Pinar del Río, zona de mayor afectación por Fasciolosis bovina, resultados que se evidencian de acuerdo a los resultados experimentales. Se demuestra, además, que el modelo ha contribuido favorablemente como herramienta de apoyo a la toma de decisiones, con el fin de cuidar la salud animal y conservar la masa ganadera.
- ✓ Los sistemas implementados en los departamentos de Sanidad Animal (Sistema Estadístico Veterinario y Sistema de Información y Vigilancia Epizootiológica), sentaron las bases para gestionar la información y analizar los datos sobre las enfermedades que mayor afectación ocasionan en la ganadería.
- ✓ La información gestionada por los sistemas referidos, ha contribuido al intercambio de experiencia de los especialistas de salud animal de las instituciones agropecuarias en la provincia y la nación. Con mayor aceptación se presenta esta situación una vez aplicado el modelo **MRB- DiagPron**, por disminuirse la complejidad para el diagnóstico de enfermedades en la ganadería, también se ha mejorado con la implementación del modelo la colaboración entre los especialistas de salud animal en materia de aplicación de medidas para disminuir la cantidad de animales enfermos y muertos por Fasciolosis bovina.

En general al aplicar el modelo se ha aumentado la concientización para emitir, almacenar y procesar la información adecuadamente. Se obtienen los diagnósticos de enfermedades ganaderas de forma rápida y certera, desde los datos gestionados y analizados. Los especialistas

de salud animal, con la aplicación del modelo, poseen información precisa para la toma de decisiones, en aras de disminuir las afectaciones que con frecuencia sufre la ganadería.

### **3.12. Factibilidad económica**

La utilización del modelo **MRB-DiagPron**, soportado por la herramienta informática **DiagPron**, permitió el ahorro de tiempo, de recursos humanos y materiales empleados diariamente para la atención veterinaria, la cual genera un costo que a partir de las características de cada enfermedad y las individualidades de cada especie varía significativamente. Por lo que el modelo influye económicamente en varios indicadores, como se muestra a continuación:

- Disminución de los tiempos de atención médica a los animales enfermos.
- Disminución de los costos por aplicación de medicamentos.
- Recuperación de la masa ganadera susceptible.

#### **Premisas para el análisis de factibilidad económica:**

Se analiza el costo por enfermedad, particularizando los análisis a partir de las características propias de cada enfermedad en cada especie. El tiempo de recuperación de la enfermedad, como principal indicador a seguir, debido a que genera efectos negativos en el sistema de salud animal como:

- ✓ Aumento en los costos
- ✓ Tratamiento inadecuado
- ✓ Riesgos de eventos adversos

En la tabla 3.7, se presenta el análisis económico del ahorro que implica la aplicación del modelo **MRB-DiagPron**. Para ello se analizan los indicadores definidos de acuerdo a los reportes de sospecha de la enfermedad de Fasciolosis bovina a través de los síntomas de ícteros,

mucosas pálidas, letargo, hepatomegalia, reportadas al Sistema de Información y Vigilancia Epizootiológica.

**Tabla 3.7.** Análisis de la implicación económica del modelo **MRB-DiagPron**. Fuente:

Elaboración propia.

<b>Indicadores</b>		<b>Sospechas de la enfermedad de Fasciolosis: Ícteros, mucosas pálidas, letargo, hepatomegalia.</b>	
<b>Costo por animal</b>		<b>Cantidad de días afectados</b>	<b>Implicación del modelo</b>
Tiempo en el proceso de atención	-	15 días	Seguimiento de la sospecha de la enfermedad dado el síntoma de Hepatomegalia aguda, controlando y mejorando los tiempos de ejecución de las actividades asistenciales para evitar la propagación a los animales sanos.
Costos de medicamentos y medios diagnósticos	\$7000 <b>\$10000</b>	-	La ocurrencia cada vez menor de intervenciones con medicamentos no satisfactorias posibilitaría la no utilización de material adicionales.
Aprovechamiento de recursos	-	-	A partir de una correcta planificación de la aplicación de medicamentos y medios de diagnóstico de acuerdo al tipo de seguimiento y complejidad de la enfermedad.
Recuperación de animales sospechosos	-	Un mes <b>45 días</b>	Si el proceso de recuperación es menor, la recuperación de los animales sospechosos será en un periodo de tiempo más corto
<b>Costo total</b>		<b>\$7000 / \$10000</b>	-



Los valores correspondientes a costos de medicamentos, medios diagnósticos y recuperación de animales sospechosos, constituyen las desviaciones promedio que se registran al existir problemas que afecten la calidad de la atención prestada a los animales enfermos, en las cuales el modelo desarrollado tiene una implicación positiva. Su implementación en el Sistema Nacional de Vigilancia Epizootiológica perteneciente al Instituto Nacional de Medicina Veterinaria, en las instituciones agropecuarias del país y en los departamentos provinciales de Sanidad Animal, no genera costo adicional, debido a que estas instituciones tienen incorporado en su presupuesto una partida para el avance tecnológico, lo cual es favorable para poder aplicar el modelo.

De acuerdo al esquema de la actividad Científico - Técnica y en particular en la prestación de los Servicios Científico Técnicos Especializados, vigente en el Centro Meteorológico Provincial de Pinar del Río, rectorado por el Instituto de Meteorología, y lo establecido por el Ministerio de Economía y Planificación (MEP), se confecciona la ficha de costo del modelo **MRB-DiagPron**. Ficha de costo que se aplica, debido a que el modelo constituye un producto, que a partir de sus resultados es posible brindar información oportuna a las instituciones agropecuarias de la provincia.

Las informaciones a las instituciones agropecuarias, van dirigidas al comportamiento de la masa ganadera, representado una herramienta útil para apoyar la toma de decisiones, disminuir afectaciones en la ganadería y cuidar la salud animal. El costo del modelo **MRB-DiagPron**, y la herramienta informática que lo instancia, según ficha de costo posee un valor es de **veinticinco mil cuatrocientos setenta y ocho pesos con treinta y seis centavos en CUP (25,478.36)**, desglosándose como se muestra en la tabla 3.8.

**Tabla 3.8.** Costo de la herramienta informática como instanciación del modelo. Fuente: Elaboración propia.

<b>FICHA DE COSTO PARA DETERMINAR EL PRECIO</b>			
<b>DESCRIPCION DEL PRODUCTO: <i>MRB-DiagPron</i></b>			
<b>Conceptos de Gastos</b>	<b>Moneda</b>	<b>Moneda</b>	<b>Moneda</b>
	<b>Total</b>	<b>Convertible</b>	<b>Nacional</b>
<b>Materias Primas y Materiales</b>	<b>22.51</b>	<b>0.00</b>	<b>22.51</b>
Materias Primas y Materiales	22.50	0.00	22.50
<b>Sub total (Gastos de Elaboración)</b>	<b>22,132.59</b>	<b>0.00</b>	<b>22,132.59</b>
<b>Otros Gastos Directos</b>	<b>21,900.10</b>	<b>0.00</b>	<b>21,900.10</b>
<b>Gastos de Fuerza de trabajo</b>	<b>228.03</b>	<b>0.00</b>	<b>228.03</b>
Salarios Devengados	200.03	0.00	200.03
Contribución a la Seguridad Social	28.00	0.00	28.00
<b>Gastos Indirectos de Producción</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>
<b>Gastos Generales y de Administración</b>	<b>4.46</b>	<b>0.00</b>	<b>4.46</b>
Energía Eléctrica	1.92	0.00	1.92
Depreciación	2.18	0.00	2.18
Otros	0.36	0.00	0.36
<b>Gastos Totales o Costo de Producción</b>	<b>22,155.10</b>	<b>0.00</b>	<b>22,155.10</b>
<b>Margen de Utilidad seg/ base autorizada</b>	<b>3,323.26</b>	<b>0.00</b>	<b>3,323.26</b>
<b>PRECIO SEGÚN LO ESTABLECIDO POR EL MEP</b>	<b>25,478.36</b>	<b>0.00</b>	<b>25,478.36</b>
<b>% Sobre el Gasto Autorizado</b>	<b>15.00</b>		<b>15.00</b>

Asimismo, el costo de las herramientas informáticas similares, a nivel mundial, se encuentran en el orden de varios millones de USD, al respecto no se tiene una referencia exacta, debido a que las empresas desarrolladoras y comercializadoras de software no proporcionan con exactitud los costos, las funcionalidades de los productos de software que poseen, por cuestiones estratégicas y de seguridad. A nivel internacional, dichas empresas poseen esquemas de negocio que solo detallan aspectos esenciales que cubren y los principales grupos de funcionalidades que proveen, a partir de lo cual se negocian los costos.

### **3.13. Implicación social**

En el orden social, el modelo desarrollado contribuye a elevar el cuidado de la salud animal y conservación de la masa ganadera, al recomendar a los especialistas de salud animal en presencia de qué enfermedad se encuentran los animales enfermos dados síntomas, signos y factores de riesgo. La herramienta informática es de utilidad práctica en la toma de decisiones, provee un mejor entendimiento y usabilidad por parte de usuarios no expertos, respecto a otras herramientas informáticas existentes para el diagnóstico de enfermedades.

### **3.14. Conclusiones del capítulo**

- La herramienta informática implementada, como instanciación del modelo **MRB-DiagPron**, permite el diagnóstico de enfermedades en la ganadería, a partir de la gestión del conocimiento, Análisis inteligente de los datos, construcción de red bayesianas y diagnóstico de enfermedades en la ganadería.
- Las indicaciones metodológicas planteadas facilitan la aplicación del modelo en redes bayesianas a partir de su instanciación en la herramienta informática.
- El diseño experimental desarrollado favoreció la evaluación estadística del **MRB-DiagPron** en casos reales. A través de diferentes pruebas estadísticas como *t-student* y análisis del

estadístico de *Shapiro-Wilk* donde se corroboró diferencias significativas del indicador de certeza de los diagnósticos por Fasciolosis bovina.

- La triangulación metodológica permitió constatar la confiabilidad de los resultados que por separado se obtuvieron a través de la escala de *Osgood*, el experimento y la entrevista a profundidad.

### CONCLUSIONES

Luego de realizada la investigación se arribó a las siguientes conclusiones:

1. A partir de la sistematización de los principales referentes teóricos que sustentan la investigación, se confirma que los modelos para el diagnóstico de enfermedades en la ganadería identificados en la literatura presentan limitaciones principalmente en relación al problema planteado en la presente investigación, fundamentándose la necesidad del desarrollo de un nuevo modelo para mejorar el diagnóstico relacionado con la Fasciolosis Bovina.
2. Constituye una necesidad en Cuba mejorar los diagnósticos de enfermedades en la ganadería y en particular de la Fasciolosis bovina y así disminuir las pérdidas que se producen producto de esa enfermedad.
3. La caracterización del procesamiento de la información en el departamento provincial de Sanidad Animal en la provincia de Pinar del Río y en las instituciones agropecuarias caso de estudio, relacionado con el diagnóstico de enfermedades en la ganadería, constituyeron la base del modelo desarrollado por la autora.
4. Teniendo en cuenta el diagnóstico realizado y con los elementos teóricos y prácticos más actuales de las ciencias informáticas se desarrolló el modelo **MRB-DiagPron** desarrollado para el diagnóstico de la Fasciolosis bovina que permite brindar información con mayor grado de certeza para el apoyo de la toma de decisiones. Su aplicación con el “Sistema de Información y Vigilancia Epizootiológica” posibilitó la automatización de los diagnósticos de enfermedades, además de constituir el instrumento de medición utilizado en la experimentación.
5. La herramienta informática implementada, como instanciación del modelo propuesto

facilita la aplicación del modelo **MRB-DiagPron** en el Sistemas de Información y Vigilancia Epizootiológica en la ganadería.

6. Los métodos científicos empleados para la validación de la propuesta de solución permitieron comprobar que los constructos del mismo son adecuados y están alineados a las tecnologías más actuales, existe una alta satisfacción de los usuarios y los experimentos realizados evidenciaron una disminución de los animales muertos por Fasciolosis bovina.

## **RECOMENDACIONES**

- ✓ Extender el modelo basado en redes bayesianas para el proceso de pronóstico de enfermedades en la ganadería, con el fin de medir la probabilidad de que ocurran determinadas situaciones en el transcurso del tiempo.
- ✓ Realizar la estimación de la probabilidad posterior de variables no conocidas en base a las variables conocidas.

**REFERENCIAS BIBLIOGRÁFICAS**

1. Alexopoulos, E., Dounias, G. D., y Vemmos, K. (1999). Medical Diagnosis of Stroke Using Inductive Machine Learning. *Machine Learning in Medical Applications*. Chania, Greece, 20-23.
2. Amaral, J. L. M., Lópes, A. J., Jansen, J. M., Faria, A. C. D., Melo, P. L. (2012). Machine learning algorithms and forced oscillation measurements applied to the automatic identification of chronic obstructive pulmonary disease. *Comput Methods Programs Biomed. Elsevier Ireland Ltd*; 2012 Mar; 105, 183-193. Recuperado de: <http://www.ncbi.nlm.nih.gov/pubmed/22018532>
3. Arias, J. P., Cardona J. F., Herrera A. C., Cerón, M. F. (2014). Rentabilidad de vacas Holstein en Antioquia: Parámetros genéticos y progreso genético. *Livestock Research for Rural Development. Volume 26, Article #121*. Retrieved June 23, 2017, from <http://www.lrrd.org/lrrd26/7/rami26121.htm>
4. Atkov O. Y, Gorokhova S. G, Sboev A. G. (2012). Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters. *J Cardiol* [Internet]. 2012/01/06 ed. 2012; 59:190- 194. Recuperado de: <http://www.ncbi.nlm.nih.gov/pubmed/22218324>
5. Barba-Romero, S. (1994), Evaluación multicriterio de proyectos, en Martínez, E. (Ed.), *Ciencia, Tecnología y Desarrollo: Interrelaciones Teóricas y Metodológicas*, (Ed. Nueva Sociedad, Caracas, pp. 455-507)
6. Barber, D. (2014). Probabilistic Modelling and Reasoning. *The Junction Tree Algorithm*. Recuperado de: <http://www.cs.helsinki.fi/u/bmmalone/probabilistic-modelspring-2014/JunctionTreeBarber.pdf>, (Consultado en mayo 2014).
7. Ben-Gal, I., Ruggeri, F., Faltin, F. y Kenett, R. (2007). Bayesian Networks. *Encyclopedia*



- of Statistics in Quality and Reliability. *John Wiley and Sons*, New York.
8. Bengoetxea, E., Larrañaga, P., Bloch, I., Perchant, A., Boeres, C. (2002). Learning and simulation of Bayesian networks applied to inexact graph matching. *Pattern Recognition*. *En prensa*.
  9. Berger, J. O. (2006). The case for objective Bayesian analysis. *Bayesian Analysis*, 1(3): 385 – 402.
  10. Bermejo, I., Oliva, J., Díez, F. J., Aries, M. (s.f). Interactive learning of Bayesian Networks using OpenMarkov.
  11. Berrios, P. (2011). Enfermedades virales de los animales domésticos. Situación en Chile. *Santiago: Chile*.
  12. Bernal, E., A. (2014). Sistema prototipo de entrenamiento pediatra para el proceso de adaptación neonatal, (Tesis de maestría en Ingeniería de Sistemas y Computación), Línea de investigación: Ingeniería del Software, Sistemas Inteligentes, *ÁREA: Sistemas Inteligentes*.
  13. Blaha, M. (2010). Patterns of data modeling. Georgia Institute of Technology, College of Computing, Atlanta, Georgia, EEUU.
  14. Blanco, R., Larrañaga, P., Inza, I. (2002). Learning Bayesian networks in the space of structures by estimation of distribution algorithms. *International Journal of Intelligent Systems*. *En prensa*.
  15. Bonet, I. (2008). Modelo para la clasificación de secuencias, en problemas de la bioinformática, usando técnicas de inteligencia artificial. (Tesis presentada en opción al grado científico de Doctor en Ciencias Técnicas). Universidad Central “Marta Abreu” de las Villas, Facultad de matemática, física y computación. Departamento de Ciencia de la Computación, Cuba.

16. Bonfa, I. (1993). HERMES: An expert system for prognosis of hepatic disease, *Proceedings of First New Zealand Internatinal Two Stream Conference on Artificial Neural Networks and Expert Systems*, pp. 240-246.
17. Bravo, F., Cruz. P. (2015). Estudios de exactitud diagnóstica: Herramientas para su Interpretación, *Revista Chilena de Radiología*. Vol. 21 N° 4, año 2015; 158-164.
18. Bright, T. J., Wong, A., Dhurjati, R., Bristow, E., Bastian, L., Coeytaux, R. R, Samsa, G., and Lobach, D. (2012). Effect of clinical decision – supporte systems a systematic review. *Annals of Internal Medicine*, 157(1): 29-43.
19. Buntine, W. (1994). Operations for learning with graphical models. *Journal of Artificial Intelligence Research*, 2, 159–225.
20. Burch, J. G, Strater F. R. (1981). *Information Systems: theory and practice*. John Wiley, New York.
21. Cabeza, (2010). Control Homeopático de Fasciolosis ovina. (Tesis en opción al grado académico de maestro en ciencias veterinarias). *Especialidad: Medicina preventiva veterinaria, Mención salud animal*. Universidad Agraria de la Habana, “Fructuoso Rodríguez Pérez”. Facultad de Medicina Veterinaria.
22. Calle, F. (2014). Técnica bayesiana de apoyo a la toma de decisiones y sus aplicaciones. (Tesis en opción al grado científico de Doctor en Ciencias Matemáticas). *Departamento de Matemática, Universidad de Extremadura, España*.
23. Castillo, E., Kjaerulff, U. (2003). Sensitivity analysis in Gaussian Bayesian networks using a symbolic-numerical technique. *Reliability Engineering and System Safety*, 79,139-148.
24. Castillo, E., Gutiérrez, J. M., Hadi, A. S. (1997). *Expert Systems and Probabilistic Network Models*, Springer-Verlag, New York.
25. Castillo, E., Kjaerulff, U., Van der Gaag, L.C. (2001). *Sensitivity Analysis in Gaussian*

- Networks. En Proceedings of Sensitivity Analysis in Model Output, *SAMO*, Madrid, España.
26. Castro, J. A., González, M., Mezo, M. (2008). Principales parasitosis en el ganado vacuno lechero: Pautas racionales de control. Laboratorio de Parasitología. *Departamento de Producción Animal, Centro de Investigaciones Agrarias de Mabegondo-Xunta de Galicia, España.*
  27. Castro, M. (2014). Imputación de datos faltantes en un modelo de tiempo de fallo acelerado. (Tesis de fin de Máster en Técnicas Estadísticas). *Universidad de Santiago de Compostela, Galicia, España,*
  28. Cerón, M. F., Herrera, R. C., Múnera, B. O., Rodríguez, Q.C., Díaz, G.A., Vergara, G.O. (2014). Manejo de información zootécnica en hatos lecheros: Evaluación genética de ganado Holstein. *Antioquia*. p.: il; 17X24 cm. ISBN: 978-958-8848-48-8.
  29. Cerón, M. F., Vasco, L. F, Rendón, J. (2013). Modelación a las ciencias animales: Generalidades de R-project. Medellín: Biogénesis, Colombia.
  30. Cleophas, T. J., Zwinderman, A. H. (2013). Machine Learning in Medicine part 2, Editorial Springer, Dordrecht Heidelberg New York. Recuperado de: <http://link.springer.com/10.1007/978-94-007-6886-4>
  31. Cobb, B. R., Rumi, R. Salmerón, A. (2007). Bayesian networks models with discrete and continuous variables. *Advances in Probabilistic Graphical Models, Studies in Fuzziness and Soft Computing. Springer*, 81-102.
  32. Chan, H. and Darwiche, A. (2004). Sensitivity analysis in Bayesian networks: From single to multiple parameters. Proceedings of the Twentieth Conference on Uncertainty in Artificial Intelligence, *AUAI Press, Virginia, Arlington*, 67-75.
  33. Charles River Analytics, I. (2004). About Bayesian Belief Networks. Charles River

- Laboratories International Recuperado de: <http://www.cra.com/commercial-solutions/beliefnetwork-modeling.asp>
34. Chávez, M. C., Rodríguez, L. O. (2002). Bayshell, Software para crear redes Bayesianas e inferir evidencias en la misma. Copyright
  35. Chávez, M. C., Casas, G., Bello, R., Grau, R. (2008a). Modelo de red bayesiana para predicción de mutaciones en secuencias de la transcriptasa inversa del VIH usando PSO. Memorias de XIV Congreso Latino-Iberoamericano en Investigación de Operaciones (CLAIO 2008): <http://socio.org.co/CLAIO2008>
  36. Chávez, M. C., Casas, G., Falcón, R., Moreira, J. L., Grau, R. (2007b). Building Fine Bayesian Networks Aided by PSO-based Feature Selection. IN: ALEXANDER, G., MORALES, K. & FERNANDO, A. (Eds.) MICA I 2007, LNAI 4827: 441- 451.
  37. Chávez, M. C., Casas, G., González, E., Grau, R. (2007a). BYNET Herramienta computacional para aprendizaje e inferencias de redes bayesianas en aplicaciones Bioinformáticas. Memorias de la XII Convención y Expo Internacional de Informática, INFORMÁTICA 2007, La Habana, Cuba, ISBN:978-959-286-002-5.
  38. Chávez, M. C., Casas, G., Moreira, J., González, E., Bello, R., Grau, R. (2008c). Uso de redes bayesianas obtenidas mediante Optimización de Enjambre de Partículas para el diagnóstico de la Hipertensión Arterial. Octavo Congreso Internacional de Investigación de Operaciones, *Revista Investigación Operacional* 30(1): 52-59.
  39. Chávez, M. C., Casas, G., Moreira, J., Silveira, P., Moya, I., Bello, R., Grau, R. (2008d). Predicción de mutaciones en secuencias de la proteína transcriptasa inversa del VIH usando nuevos métodos para Aprendizaje Estructural de Redes Bayesianas. *Revista Avances en Sistemas e Informática* 4(2): 77-85.
  40. Chávez, M. C., Casas, G., Moya, I., Grau, R. (2008b). A new Method for Learning

- Bayesian Networks. Application to Data *Splice* Site Classification. Proceedings of the Second Workshop on Bioinformatics Cuba, Flanders *IWOBI 2008, Santa Clara, Cuba*, ISBN:978-959-250-394-6.
41. Chávez, M. C., Gladys Casas, G., Moreira, J. E., Bello, R., Grau, R. (2008). Predicción de Mutaciones en Secuencias de la Proteína Transcriptasa Inversa del VIH usando nuevos Métodos para Aprendizaje Estructural de Redes Bayesianas, *Revista Avances en Sistemas e Informática, Medellín*, ISSN - 16577663, 5, 2.
  42. Chávez, M. C., Grau, R., García, M. M. (1999). Un método para construir Redes Bayesianas. *Revista Facultad de Ingeniería* 19: 76-84
  43. Chávez, M. C., Grau, R., Sánchez, R. (2005). Construcción de árboles filogenéticos a partir de secuencias de ADN y su integración en una red bayesiana. *Memorias de la XI Convención Expo Internacional de Informática, INFORMÁTICA 2005, La Habana, Cuba* ISBN: 978-959-716-487-6.
  44. Chávez, M. C., Grau, R., Casas, G. (2008). Las redes bayesianas y la bioinformática. En: Modelos de redes bayesianas en el estudio de secuencias Genómicas y otros problemas biomédicos. *Universidad Central "Marta Abreu" de las Villas, Facultad de Matemática, Física y Computación, Villa Clara, Cuba*, 11- 115.
  45. Chávez, M. C., Silveira, P., Casas, G., Grau, R., Bello, R. (2007). Aprendizaje estructural de redes bayesianas utilizando PSO. *Memorias en Boletín de la Sociedad Cubana de Matemática*, 5, Trabajo IA7, Número Especial en CD de COMPUMAT, ISSN: 1728-6042, Holguín, Cuba.
  46. Chávez, M. C., Silveira, P., Casas, G., Grau, R., Bello, R. (2007c). Aprendizaje estructural de redes bayesianas utilizando PSO. *Memorias en Boletín de la Sociedad Cubana de Matemática*, Trabajo IA7, Número Especial en CD de COMPUMAT, Holguín, Cuba 5.

47. Chow, C., Liu, C. (1968). Approximating discrete probability distributions with dependence trees. *Information Theory, IEEE Transactions on*, 14(3), 462-467.
48. Chuang C. L. (2011). Case-based reasoning support for liver disease diagnosis. Recuperado de: <http://www.ncbi.nlm.nih.gov/pubmed/21757326>
49. Daly, R., Shen, Q., Aitken, S. (2011). Review: Learning Bayesian networks: Approaches and issues, *The knowledge engineering review*, 26(2), 99-157.
50. Darmoni, S. J, Massari, P., Droy, J. M., Thierry Blanc, N. M., Moiro, E., Leroy, J. (1994). SETH: an expert system for the management on acute drug poisoning in adults, *Computer Methods and Programs in Biomedicine, Elsevier*. Recuperado de: [https://doi.org/10.1016/0169-2607\(94\)90067-1](https://doi.org/10.1016/0169-2607(94)90067-1)
51. Daut E. F, Lahodny G. J., Peterson M., J., Ivanek R. (2016). Interacting Effects of Newcastle Disease Transmission and Illegal Trade on a Wild Population of White-Winged Parakeets in Peru: *A Modeling Approach. PLoS ONE*, Vol. 11.
52. Díaz A., Torres V., Herrera S. M., Fernández L., Sarduy L. (2014). Modelación del crecimiento de bovinos en pastoreo con gramíneas y leguminosas. *Zootecnia Tropical*, 32(4), 363-376.
53. Dietzel, M., Baltzer, P. A, Dietzel A. (2011). Artificial Neural Networks for differential diagnosis of breast lesions in MR-Mammography: A systematic approach addressing the influence of network architecture on diagnostic performance using a large clinical database. *Eur J Radiol [Internet]*. 2011/04/05 ed. 2011. Recuperado de: <http://www.ncbi.nlm.nih.gov/pubmed/21459533>
54. Diez, F. J. (2010). Introducción a los Modelos Gráficos Probabilistas. *Departamento de Ingeniería Artificial. UNED*.

55. Díez, J. F. (1998). Aplicaciones de los modelos gráficos probabilistas en medicina. En: J. A. Gámez y J. M. Puerta (eds.), *Sistemas Expertos Probabilísticos*, *Universidad de Castilla-La Mancha*, 239-263.
56. Doldán, F. (2007). Redes Bayesianas y Riesgo Operacional. *Revista Gallega de Economía* 16 (Número extraordinario): [http://www.usc.es/econo/RGE/Vol16\\_ex/Castelan/art1c.pdf](http://www.usc.es/econo/RGE/Vol16_ex/Castelan/art1c.pdf)
57. Driver, E., Morrell, D. (1995). Implementation of continuous bayesian networks using sums of weighted Gaussians. *In Proceedings of the Eleventh UAI Conference*, 134- 140. *Morgan Kaufmann*.
58. Druzdal, M. J. (1993). Probabilistic reasoning in decision support systems: from computation to common sense. (PhD thesis), Carnegie Mellon University.
59. Edwards, W. (1998). Tools for and experiences with Bayesian normative modeling, *American Psychologist*, 53(4).
60. Febles, O. (2012). Modelo para el desarrollo de aplicaciones compuestas basadas en arquitecturas orientadas a servicios. (Tesis en opción al grado científico de Doctor en Ciencias Técnicas, especialidad- Informática). Universidad de las Ciencias Informáticas. Ciudad de la Habana, Cuba.
61. Font, E., Lezcano, C., Palenque, E., Estrada, V., Febles, J. P. (2014). Gestión de la Información y el Conocimiento. *Editorial Universitaria, Félix Varela*, 1-211, ISBN: 976-959-07-1805-2.
62. Franco, A. M., (2014). Modelo para Análisis de Riesgo de la Diabetes Mellitus 2, usando Inteligencia de Negocios y Minería de Datos. (Tesis para optar al título de Magister en Ingeniería de Sistemas y Computación). *Universidad Nacional de Colombia, Facultad de Ingeniería, Departamento de Sistemas e Industrial*, Bogotá, D.C.
63. Gambelli, D., Bruschi, V. (2010). A Bayesian network to predict the probability of organic

- farms exists from the sector: A case study from Marche, Italy. *Computers and Electronics in Agriculture*, 71, 22-31.
64. Gámez, J. A., Puerta, J. M. (1998). Sistemas expertos probabilísticos, *Univ de Castilla La Mancha, volumen 20*.
65. García, F., Calderón, A., Gallego, J. F., Jiménez, G., Márquez, D., Arango, D., Arreaza. L. C. (2011). Másleche 2.0. *Revista Corpoica*, 12, ISSN: 0122-8706.
66. García, L. (1990). Probabilidad e Inteligencia Artificial. *Conferencias de Laureano García*, Universidad de la Habana, Cuba.
67. García, Ó. (2015). Futuro de la enseñanza médica: inteligencia artificial y big data. *FEM: Revista de la Fundación Educación Médica*, 18 (Supl. 1), s60-s61.  
<https://dx.doi.org/10.4321/S2014-98322015000300009>
68. García, Y. (2011). Trabajo investigativo sobre Minería de Datos. Palacio Provincial de la Computación. Santiago de Cuba.
69. Geenen P. L., Gaag L. C., Loeffen W. L. A., Elbers A. R. W. (2011). Constructing naive Bayesian classifiers for veterinary medicine: A case study in the clinical diagnosis of classical swine fever. *Res Vet Sci [Internet]. Elsevier Ltd*; 2011; 91:64-70. Disponible en: <http://dx.doi.org/10.1016/j.rvsc.2010.08.006>
70. Geenen, P. L., Van der Gaag, L. C. (2005). Developing a Bayesian network for clinical diagnosis in veterinary medicine: from the individual to the herd. In Proceedings of the Third Bayesian Modelling Applications Workshop: *Edinburgh*.
71. Getoor, L., Rhee, J. T., Koller, D., Small, P. (2004). Understanding tuberculosis epidemiology using structured statistical models. *Artificial Intelligence in Medicine*, 30, 233-256.
72. Gilk, W. R., Richardso, S., Spiegelhalter, D. J. (1996). *Markov Monte Carlo in practice*,



- Chapman and Hall.
73. Gómez, E., Gómez-Villegas, M. A. Marín, J. M. (1998). A multivariate generalization of the power exponential family of distributions. *Communications in Statistics-Theory and Methods*, 27:3, 589-600.
  74. González, N., Estrada, V., Romillo, A. (2016). Herramienta en un entorno web para el diagnóstico y pronóstico de enfermedades en la ganadería, *Revista Cubana de Ciencias Informáticas Vol. 10, No. 4, octubre - diciembre, 2016, ISSN: 2227-1899 | RNPS: 2301 <http://rcci.uci.cu>, Pág. 112-124.*
  75. González, N., Estrada, V., Romillo, A., Quintero, A. (2017). Estudio de Caso. Diagnóstico y pronóstico de la Fasciolosis bovina basado en redes bayesianas, *REDVET - Revista electrónica de Veterinaria Volumen 18 N° 9 - - ISSN 1695-7504. <http://www.veterinaria.org/revistas/redvet/n090917.html>,*
  76. González, C., Lozano, J. A., Larrañaga, P. (2002). Mathematical modelling of UMDAc algorithm with tournament selection. Behavior on linear and quadratic functions. *International Journal of Approximate Reasoning*. En prensa.
  77. González, C., Lozano, J. A., Larrañaga, P. (2000). Analyzing the PBIL Algorithm by Means of Discrete Dynamical Systems, *Complex Systems*, 12(4), 465–479.
  78. González, N. (2014). Modelo matemático para el diagnóstico y pronóstico de Fasciola Hepática en el ganado bovino, *Volumen 15 N° 11 - <http://www.veterinaria.org/revistas/redvet/n111114.html>, REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504.*
  79. González, N., Miranda, C., A., Malagón, G., M. (2014). Gestión organizacional para el desarrollo local a través de un modelado matemático en Redes Bayesianas aplicado en la Medicina Veterinaria, *Revista Avances Vol. 16 (1), ene. -marzo., 2014.*

80. Grau, R., Correa, C. y Rojas, M. (2004). Metodología de la Investigación Segunda Edición, *EL POIRA Editores S.A.*, Ibagué, Colombia, ISBN: 958-8028-10-8.
81. Grau, R., Chávez, M. C., Sánchez, R., Morgado, E., Casas, G. y Bonet, I. (2007a). Boolean algebraic structures of the genetic code. Possibilities of applications. Lecture Notes on Bioinformatics, Knowledge Discovery and Emergent Complexity in Bioinformatics 4366: 10-21.
82. Grau, R., Chávez, M. C., Sánchez, R., Morgado, E., Casas, G. y Bonet, I. (2007b). Boolean algebraic structures of the genetic code. Possibilities of applications. IN: TUYLS, K. *et al.* (Eds.). KDEB 2006, LNBI 4366: 10–21.
83. Gutiérrez, J. M., Cano, R., Sordo, C. (2005). Analysis and downscaling multi-model seasonal forecasts in Peru using self-organizing maps. *Tellus*, 57A,435–447.
84. Gutiérrez, J. M., Cofino, A. S., Cano, R., Rodríguez, M. A. (2004). Clustering methods for statistical downscaling in short-range weather forecast. *Rev. Mon. Wea.*, 132(9), 2169 - 2183.
85. Hand, D. J., Mannila, H. and Smyth, P. (2001). Principles of data mining. *MIT press*, 53.
86. Heckerman, D. (1996). A Tutorial on Learning with Bayesian Networks. Microsoft Research Tech. Report MSR-TR-95-06, Redmond, WA:  
<ftp://ftp.research.microsoft.com/pub/dtg/david/tutorial.ps>
87. Heckerman, D., Geiger, D. (1995). Learning Bayesian networks: unification for discrete and Gaussian domains. In P. Besnard and S. Hanks (Eds.), *San Francisco, Calif.: Morgan Kaufmann, Proc. Eleventh Conference on Uncertainty in Artificial Intelligence (UAI '95)*, 274–284.
88. Heckerman, D., Geiger, D., Chickering, D. M. (1995). Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, 20, 197–243.

89. Heckerman, D., Meek, C., Cooper, G. (1997). A Bayesian approach to causal discovery. *Microsoft Research. Technical Report MSR-TR-97-05.*
90. Heckerman, D., Shachter, R. (1995). Decision theoretic foundations for causal reasoning. *Journal of A.I. Research*, 3, 405 - 430.
91. Hernández, R. (2008). Metodología de la investigación. *Edition ed.* La Habana.
92. Hernández-Sampieri R., Fernández C., Baptista P. (1998). Metodología de la Investigación. Segunda Edición, Cámara Nacional de la Industria Editorial Mexicana, Reg. No. 736. 1998.
93. Højsgaard, S., Edwards, D., Lauritzen, S. (2012). Graphical Models with R, *Springer*, USA.
94. Huang, M. L, Hung Y. H, Lee, W. M., Li, R. K, Wang, T.H. (2010). Usage of case-based reasoning, neural network and adaptive neuro-fuzzy inference system classification techniques in breast cancer dataset classification diagnosis. *J Med Syst [Internet]. 2010/08/13 ed. 2012; 36:407-414. Recuperado de: <http://www.ncbi.nlm.nih.gov/pubmed/20703710>*
95. Hussein, A. (2015). The use of triangulation in social sciences research: Can qualitative and quantitative methods be combined? *Journal of Comparative Social Work*, 4(1).
96. Inza, I., Larrañaga, P., Etxeberria, R., Sierra, B. (2000). Feature Subset Selection by Bayesian networks based optimization. *Artificial Intelligence*, 123,157–184.
97. Inza, I., Merino, M., Larrañaga, P., Quiroga, J., Sierra, B., Giralda M. (2001). Feature subset selection by genetic algorithms and estimation of distribution algorithms. A case study in the survival of cirrhotic patients treated with TIPS. *Artificial Intelligence in Medicine*, 23–2, 187–205.
98. Inza, I., Sierra, B., Blanco, R., Larrañaga, P. (2002). Gene selection by sequential search wrapper approaches in microarray cancer class prediction. *Journal of Intelligent and*

- Fuzzy Systems*. En prensa.
99. Jones, K. (2016.) Modelo informático que predice brotes de enfermedades zoonóticas. Informática en la predicción de propagación de enfermedades animales, *departamento de Diversidad Biológica e Investigación del Medio Ambiente*. University College London (UCL) en el Reino Unido.
100. Jun, H. (2010). "Evaluating Node Importance with Multi-Criteria," in *Green Computing and Communications (GreenCom), 2010 IEEE/ACM Int'l Conference on & Int'l Conference on Cyber, Physical and Social Computing (CPSCoM), 2010*, pp. 792-797. *et al.*, "Evaluating Node Importance with Multi-Criteria," in *Green Computing and Communications (GreenCom), 2010 IEEE/ACM Int'l Conference on & Int'l Conference on Cyber, Physical and Social Computing (CPSCoM), 2010*, pp. 792-797.
101. KDnuggets, (2008). Bayesian Networks and Bayesian Classifier Software. <http://www.kdnuggets.com/software/bayesian.html>
102. Kjaerulff, U. B, Madsen, A. L. (2013). Bayesian Networks and Influence diagrams: a guide to construction and analysis, *New York: Springer*. 22.
103. Kjaerulff, U., Madsen, A. (2007). Probabilistic Networks for Practitioners. A Guide to Construction and Analysis of Bayesian Networks and Influence Diagrams. *New York: Springer*.
104. Korb, K., Nicholson, A. (2004). Bayesian Artificial Intelligence. *Boca Ratón, FL: Chapman and Hall*.
105. Lacave, C. (2002). Explicación en redes bayesianas causales. Aplicaciones médicas. (Tesis en opción al grado científico de Doctor en Ciencias Informáticas). *Departamento de Inteligencia Artificial*, Universidad Nacional de Educación a Distancia, España.

106. Lacave, C., Luque, M., & Díez, F. J. (2007). Explanation of Bayesian networks and influence diagrams in Elvira. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 37(4), 952-965.
107. Larrañaga, P., Etxeberria, R., Lozano, J. A., Peña, J. M. (2000). Combinatorial optimization by learning and simulation of Bayesian networks. *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence*, 343–352
108. Larrañaga, P., Kuijpers, C., Murga, R., Inza, I., Dizdarevich, S. (1999). Evolutionary algorithms for the travelling salesman problem: A review of representations and operators. *Artificial Intelligence Review*, 13,129–170
109. Larrañaga, P., Kuijpers, C., Murga, R., Yurramendi, Y. (1996). Learning Bayesian network structures by searching for the best ordering with genetic algorithms. *IEEE Transactions on System, Man and Cybernetics*, 26 (4), 487-493.
110. Larrañaga, P., Kuijpers, C., Poza, M., Murga, R. (1997). Decomposing Bayesian networks by genetic algorithms. *Statistics and Computing*, 7 (1), 19-34.
111. Larrañaga, P., Lozano, J. A. (2001) (eds.). Estimation of Distribution Algorithms. A New Tool for Evolutionary Computation. *Kluwer Academic Publishers*.
112. Larrañaga, P., Poza, M., Yurramendi, Y., Murga, R., Kuijpers, C. (1996). Structure learning of Bayesian networks by genetic algorithms: A performance analysis of control parameters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18 (9), 912-926.
113. Laskey, K. B. (1995). Sensitivity analysis for probability assessments in Bayesian networks. *IEEE Trans. Syst., Man, Cybern*, 25, 901-909.
114. Lauritzen, S. (1992). Propagation of probabilities, means and variances in mixed graphical association models. *Journal of the American Statistical Association*, (87),1098-1108.

115. Lauritzen, S. L., Jensen, F. (2001). Stable local computation with conditional Gaussian distributions. *Statistics and Computing*, 11, 191–203.
116. Lewis, F. I., Brülisauer, F., Gunn, G. J. (2011). Structure discovery in Bayesian networks: An analytical tool for analyzing complex animal health data. *Preventive veterinary medicine*, 100, (2),109-115.
117. Leyva M. Y. (2013). Modelo de ayuda a la toma de decisiones basado en mapas Cognitivos Difusos. (Tesis en opción al grado científico de Doctor en Ciencias Técnicas). Universidad de las Ciencias Informáticas. Ciudad de la Habana, Cuba.
118. López, J., García, J. (2010). Modelado Causal Predictivo Basado en Redes. Recuperado de: [http://www.ceautomatica.es/old/actividades/jornadas/XXVIII/documentos/1381-Model\\_causal\\_pred2.pdf](http://www.ceautomatica.es/old/actividades/jornadas/XXVIII/documentos/1381-Model_causal_pred2.pdf)
119. Lozano, J. A., Larrañaga, P. (1998). Applying genetic algorithms to search for the best hierarchical clustering of a dataset. *Pattern Recognition Letters*, 20, 911-918.
120. Lozano, J. A., Larrañaga, P., Graña, M., Albizuri F. X. (1999). Genetic algorithms: bridging the convergence gap. *Theoretical Computer Science*, 229, 11-22.
121. Lozano, M. R. (2011). El papel de las redes bayesianas en la toma de decisiones, 2. Recuperado de: [http://www.urosario.edu.co/Administracion/documentos/investigacion/laboratorio/miller\\_2\\_3.pdf](http://www.urosario.edu.co/Administracion/documentos/investigacion/laboratorio/miller_2_3.pdf)
122. Lugo, S. O., Maldonado, G., Murata, Ch. (2014). Inteligencia artificial para asistir el diagnóstico clínico en medicina, *Revista Alergia México*, 61, (2), abril - junio. *Laboratories International* Recuperado de: <http://www.cra.com/commercial-solutions/beliefnetworkmodeling.asp>
123. Mahabala, H. N., Chandrasekhara, M. K., Baskar, S., Ramesh, S., Somasundaram, M.S

- (1992). ICHT: An Intelligent Referral System for Primary Child Health Care. *Conference of the South East Asia Regional Computer Confederation. Kuala Lumpur. Proceedings SEARCC, 92, XI.*
124. Main, P., Navarro, H. (2009). Analyzing the effect of introducing a kurtosis parameter in Gaussian Bayesian networks. *Reliability Engineering and System Safety, 94, 922-926.*
125. Manickam, S., Abidi, S. S. R. (1999). Experienced Based Medical Diagnostics System Over the World Wide Web (WWW), *Proceedings of The First National Conference on Artificial Intelligence Application in Industry, Kuala Lumpur, 47 - 56.*
126. Marasco, J., Doerfler, R. (2011). Doc, what are my chances? *Artículo, The UMAP Journal.*
127. Marcos, A. P., Colón, J. Z., Gutiérrez, M. R. y Santos, A. M. P. (2014). Investigación cualitativa. *Elsevier España.*
128. Martínez E., Escudey M., (1998). Evaluación y Decisión Multicriterio. Reflexiones y Experiencias. *Editorial Universidad de Santiago/UNESCO, Santiago de Chile.*
129. Martínez, R. (2014). Metodologías basadas en minería de datos para el diseño y optimización de técnicas de clasificación automática. (Tesis en opción al grado científico de Doctor en Ciencias Informáticas). *Facultad de Informática, Universidad de Murcia, España.*
130. Matsumoto, A., Carvalho, R. N., Ladeira, M., Costa, P., Santos, L., Silva, D., Machado, E. (2011). UnBBayes: a java framework for probabilistic models in AI. *Java in Academia and Research. Java in Academia and Research, 34.*
131. McKendrick, I. J., Gettinby, G., Gu, Y., Reid, S. W. J., Revie, C. W. (2000). Using a Bayesian belief network to aid differential diagnosis of tropical bovine diseases. *Preventive Veterinary Medicine, 47, (3), 141-156.*

132. Monsalve, A., E. (2017). Sistema de ayuda a la decisión clínica en enfermedades de diagnóstico complejo. Tesis en opción al grado científico de Doctor en Ciencias Informáticas. *Departamento de Tecnología, Informática y Computación, Escuela politecnica superior de Alicante, España*. <http://www.eltallerdigital.com>
133. Murphy, K. (2005). Software Packages for Graphical Models / Bayesian Networks. <http://cs.berkeley.edu/~murphyk/Bayes/bnsoft.html>
134. Nadkarni, S., Shenoy, P. P. (2001). A Bayesian network approach to making inferences in causal maps. *European Journal of Operational Research*, 128(3), 479–498.
135. Nadkarni, S., Shenoy, P. P. (2004). A causal mapping approach to constructing Bayesian networks. *Decision support systems*, 38(2), 259-281.
136. Nagarajan, R., Scutari, M., Lèbre, S. (2013). *Bayesian Networks in R with Applications in Systems Biology*, Springer, USA, 2013.
137. Neapolitan, R. (2003). Learning Bayesian Networks. *Prentice Hall Series in Artificial Intelligence*, USA.
138. Ochoa, A., (2015). Diagnóstico para Síndrome de Asperger utilizando una aplicación inteligente a partir del razonamiento basado en casos, *Revista Electrónica de Divulgación de la Investigación Vol. 10, Diciembre – 2015*. Recuperado de: <https://www.researchgate.net/publication/291698853>
139. Ordoñez, P., Guamán, J, Patiño, S. Torres, C. (2016). Diagnóstico de enfermedades infecciosas virales, *Carrera de Ingeniería en Sistemas. Universidad Nacional de Loja*.
140. Ortiz, A, Munilla J, Górriz, J.M, Ramírez, J. (2015). Nuevo método para el diagnóstico de la enfermedad de Alzheimer basado en inteligencia artificial, *International Journal of Neural Systems*, November 2016, Vol. 26, No. 07. Recuperado de: <https://doi.org/10.1142/S0129065716500258>



141. Otto, L. Kristensen, C. S. (2004). A biological network describing infection with *Mycoplasma Hyponeumoniae* in swine heard. *Preventive Veterinary Medicine*, 66 (1), 141-161.
142. Ovalle, C. (2015). Caracterización de un modelo celular diseñado para el tamizaje genético mediante RNA de interferencia en el complejo macrófago – infectado por *Leishmania (Viannia) braziliensis*. (Tesis de en opción al grado científico de Doctor en Ciencias), *Universidad Nacional de Colombia*.
143. Passold, F. (1995). Sistema Especialista Híbrido en Anestesiología para Pacientes Críticos/Problemáticos. *M. Eng. dissertation* of Federal University of Santa Catarina, Dept. of Electrical Engineering, 1995.
144. Patel, V. L., Shortliffe, E. H., Stefanelli, M., Szolovits, P., Berthold, M. R., Bellazzi, R. (2008). The coming of age of artificial intelligence in medicine. *Artif Intell Med* [Internet]. 2008/09/16 ed. 2009; 46:5-17. Recuperado de: <http://www.ncbi.nlm.nih.gov/pubmed/18790621>
145. Pearl, J. (2000). *Causality: Models, reasoning and inference*, Cambridge, Univ. Press, 29.
146. Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: *Morgan Kaufmann*.
147. Pearl, J. (2001). Bayesian networks, causal inference and knowledge discovery. *UCLA Cognitive Systems Laboratory, Technical Report*.
148. Pearl, J. L. (1993). Graphical Models, Causality and Intervention. *Stat. Sci.* 8(3): 266- 273.
149. Peña, J. M., Lozano, J. A., Larrañaga., P. (2002). Learning recursive Bayesian multinets for clustering by means of constructive induction. *Machine Learning*, 47, 63–89.

150. Pérez, K., Leyva, M., Espinilla, M., Estrada, V. (2014). Computación con palabras en la toma de decisiones mediante mapas cognitivos difusos. *Revista Cubana de Ciencias Informáticas*, 8(2), 19-34, ISSN: 1994-1536 | RNPS: 0547, Recuperado de: <http://www.rcci.uci.cu>
151. Piñero, P. Y. (2005). Un modelo para el aprendizaje y la clasificación automática basado en técnicas de *Softcomputing*. (Tesis presentada en opción al grado de Doctor en Ciencias Técnicas), *Universidad de Ciencias Informáticas*, La Habana, Cuba.
152. Pourret, O., Onísco, (2008). *A. Bayesian Networks: A practical Guide to Applications*, Wiley, 2008.
153. Puga, J. L., García, J. G., Sánchez, L. d., Solana, E. I. (2007). Las redes bayesianas como herramientas de modelado en psicología. *Anales de psicología*, 23(2), 307-316.
154. Ramesh, A.N., Kambhampati, C., Monson, J.R., Drew, P. J. (2004). Artificial intelligence in medicine. *Ann R Coll Surg Engl* [Internet]. 2004/08/31 ed. 2004; 86:334-338. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/15333167>
155. Ramírez, J. F. (2016). Modelo para la selección de equipos de trabajo quirúrgico en sistemas de información en salud aplicando técnicas de inteligencia organizacional. (Tesis presentada en opción al grado científico de Doctor en Ciencias Técnicas), *Universidad de las Ciencias Informáticas*, La Habana, Cuba.
156. Rebane, G., Pearl, J. (2013). The recovery of causal poly-trees from statistical data. *arXiv preprint arXiv*, 1304 - 2736.
157. Regalado, R. F. (2009). El teorema de Bayes y su utilización en la interpretación de las pruebas diagnósticas en el laboratorio clínico. *Revista Cubana de Investigaciones Biomédicas*, 28(3), 158-165.
158. Rivera, M. (2011). El papel de las redes bayesianas en la toma de decisiones, *Universidad*

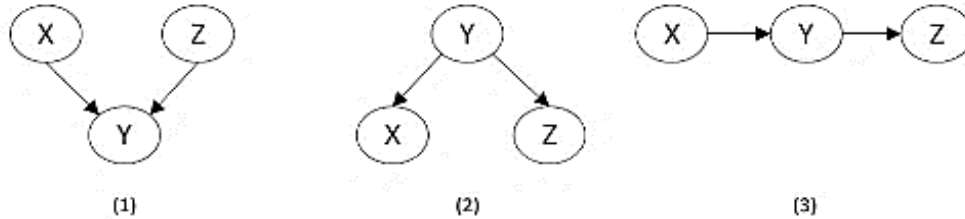
- del Rosario, *Laboratorio de Modelamiento y Simulación, Facultad de Administración. La Simulación al Servicio de la Academia*, 2, ISSN 2027 – 7709.
159. Rodríguez, D., Dolado, J. (2010). Redes Bayesianas en la Ingeniería del Software. *Universidad de Alcalá, España*.
160. Ropero, R. F., Aguilera, P. A., Fernández, A., Rumí, R. (2014). Redes bayesianas: una herramienta probabilística en los modelos de distribución de especies. *REVISTA CIENTÍFICA DE ECOLOGÍA Y MEDIO AMBIENTE*, ISSN 1697-2473/ Open Access, Recuperado de: [www.revistaecosistemas.net](http://www.revistaecosistemas.net)
161. Ruseckaite, R. (1999). Computer Interactive System for Ascertainment of Visual Perception Disorders. *Machine Learning and Applications: Machine Learning in Medical Applications. Chania, Greece, 27-29*.
162. Sabiote, C. R., Llorente, T. P., Pérez, J. G. (2006). La triangulación analítica como recurso para la validación de estudios de encuestas recurrentes de investigaciones de réplica en Educación Superior. *RELIEVE*, 12.
163. Sánchez, L., Cabañas, G., Abad, Y., Torres, V. (2014). Utilización de modelos ARIMA para la predicción de la producción de leche. Estudio de caso en la UBPC “Maniabo”, Las Tunas. *Revista Cubana de Ciencia Agrícola*, 48, 3.
164. Shahbaba, Babak, (2012). *Biostatistics with R: An introduction to Statistics Through Biological Data (Use R!)*, Springer, USA.
165. Shull, F., Singer, J., Sjoberg, D. I. K. (2008). *Guide to Advanced Empirical Software Engineering*. Edtion ed.: Springer-Verlag, p. 394, ISBN 978-1-84800-043-8.
166. Sierra, B., Serrano, N., Larrañaga, P., Plasencia, E. J., Inza, I., Jiménez, J. J., De la Rosa, J. M., Mora, M. J. (2001). Using Bayesian networks in the construction of a bi-level multiclassifier. A case study using intensive care unit patient’s data. *Artificial Intelligence*

- in Medicine*, 22, 233–248.
167. Soler, F. (2014). Estimación de sucesos pocos probables mediante Redes Bayesianas. (Tesis en opción al grado científico de Doctor en Ciencias Informáticas). *Departamento de Tecnologías y Sistemas de la Información, Universidad de Castilla La Mancha, Ciudad Real, España.*
168. Spirtes, P., Glymour, C., Sheines, R. (1993). Causation, Prediction and Search, Springer Verlag, New York.
169. Stone, J. (2013). Bayes Rule a Tutorial Introduction to Bayesian Analysis, *Sebtel Press*, England, 2013.
170. Stuart, J. R. y Norvig, N. (2003). Artificial Intelligence: A Modern Approach. *Prentice Hall; 2 edition*: 1132 páginas.
171. Susi, R. (2007). Análisis de sensibilidad en redes bayesianas Gaussianas. (Tesis en opción al grado de doctor en Ciencias Matemáticas), *Facultad de Ciencias Matemáticas. Departamento de Estadística e Investigación Operativa. Universidad Complutense de Madrid*, ISBN 978-84-669-3133-5, España.
172. Teach, R.L, Shortliffe, E.H. (1984). An analysis of physician attitudes regarding computer based clinical consultation systems. En: B. G. Buchanan y E. H. Shortliffe (eds.), *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, cap. 34, págs. 635 - 652. Addison-Wesley, Reading, MA.
173. Tenorio, J. M, Hummel, A. D, Cohrs, F. M, Sdepanian, V. L, Pisa, I. T. (2011). Artificial intelligence techniques applied to the development of a decision-support system for diagnosing celiac disease. *Int J Med Inf [Internet]*. 2011/09/16 ed. 2011;80(11):793-802. Recuperado de: <http://www.ncbi.nlm.nih.gov/pubmed/21917512>
174. Tonda, A.P., Lutton, E., Reuillon, R., Squillero, G., Wuillemin, P. (2012). Bayesian

- network structure learning from datasets through graph evolution. In *Genetic Programming, Springer*, 254-265.
175. Valencia, M. M. A. (2013). La triangulación metodológica: sus principios, alcances y limitaciones. *Investigación y educación en enfermería*, 18(1).
176. Valles, M. S. (2014). Entrevistas cualitativas (Vol. 32). *CIS-Centro de Investigaciones Sociológicas*. Madrid. ISBN: 978-84-7476-655-4.
177. Viviani, P. (2014). Análisis de sensibilidad a la evidencia en Redes Bayesianas Gaussianas, (Tesis en opción al grado de doctor en ciencias matemáticas). *Universidad Complutense de Madrid. Facultad de ciencias matemáticas. Departamento de Estadística e Investigación Operativa I*.
178. Voinov, A., Bousquet, F. (2010). Modelling with stakeholders. *Environmental Modelling and Software*, 25, 1268-1281.
179. Watson, I. (1999). Case-based reasoning is a methodology not a technology.;12:303-308
180. Wikipedia, Biografía Judea Pearl (2013). Recuperado de: [http://es.wikipedia.org/wiki/Judea\\_Pearl](http://es.wikipedia.org/wiki/Judea_Pearl), (Consultado en mayo 2014).
181. Wikipedia, Biografía Thomas Bayes (2013). Recuperado de: [http://es.wikipedia.org/wiki/Thomas\\_Bayes](http://es.wikipedia.org/wiki/Thomas_Bayes), (Consultado en mayo 2014).
182. Wikipedia, R (Lenguaje de programación) (2013). Recuperado de: <http://es.wikipedia.org/wiki/R> (lenguaje de programaci%C3%B3n) (Consultado en mayo 2014).
183. Wikipedia, Red Bayesiana (2013). Recuperado de: [http://es.wikipedia.org/wiki/Red\\_bayesiana](http://es.wikipedia.org/wiki/Red_bayesiana) (Consultado en mayo 2014).
184. Wiltaker, J. (1990). *Graphical Models in Applied Multivariate Statistical*. *Wiley Series in Probability & Statistics*: 462 páginas.

185. Zelic, I., Lavrac, N., Najdenov, P., Rener-Primec, Z. (1999). Impact of machine learning of the Diagnosis and Prognosis of First Cerebral Paroxysm. Machine Learning and Applications: *Machine Learning in Medical Applications*. Chania, Greece, 24-26.
186. Zumaquero, J. L., Sarracent, J., Rojas, R., Rojas, L., Martínez, Y. (2013). Fascioliasis and Intestinal Parasitoses Affecting Schoolchildren in Atlixco, Puebla State, Mexico: *Epidemiology and Treatment with Nitazoxanide*. *PLoS Negl Trop*, 7(11), e2553. doi: 10.1371/journal.pntd.0002553.

## ANEXOS

**Anexo 1.** Tipos de conexiones en un grafo dirigido

**Figura 1.** Conexión convergente **Figura 2.** Conexión divergente **Figura 3.** Conexión secuencial

**Anexo 2.** Conjunto de pasos lógicos para la propagación de probabilidades en poliárboles

1. Obtener el esqueleto utilizando el algoritmo de Chow y Liu.
2. Recorrer la red hasta encontrar una tripleta de nodos que sean convergentes (tercer caso) – nodo multipadre.
3. A partir de un nodo multipadre determinar las direcciones de los arcos utilizando la prueba de tripletas hasta donde sea posible (base causal).
4. Repetir (2) a (3) hasta que ya no se puedan descubrir más direcciones.
5. Si quedan arcos sin direccionar utilizar semántica externa para obtener su dirección.

**Anexo 3.** Comparación de paquetes de software de Modelos Gráficos: Redes bayesianas

<b>Nombre</b>	<b>Sistema operativo que lo soporta</b>	<b>Características</b>
<b>AgenaRisk</b>	Windows, Unix	Realiza simulación por discretización Dinámica
<b>Analytica</b>	Windows, Mac	El proceso de propagación que ejecuta es compatible con otras herramientas basadas en RB
<b>Banjo</b>	Windows, Unix, Mac	Utilizado para crear estructuras de aprendizaje dinámica o discreta, compatible con las redes bayesianas de variables discretas
<b>Bayda</b>	Windows, Unix, Mac	Clasificador Naïve bayes
<b>BayesiaLab</b>	Windows	Aprendizaje estructural, modelos dinámicos
<b>Bayesware Discoverer</b>	Windows, Unix, Mac	Usa límite y compresión para aprendizaje con datos ausentes
<b>BNT</b>	Windows, Unix, Mac	Maneja modelos dinámicos como Modelos ocultos de Markov y filtros Kalman
<b>Causal discoverer</b>	Windows	Estructura de aprendizaje
<b>CoCo+Xlisp</b>	Unix	Diseñado para tablas de contingencia
<b>GDAGsim</b>	Windows, Unix, Mac	Análisis bayesiano de modelos dirigidos gaussianos
<b>GMRFsim</b>	Windows, Unix, Mac	Análisis bayesiano de modelos no dirigidos gaussianos
<b>GMTk</b>	Unix	Diseñado para reconocimiento del habla
<b>KBaseAI</b>	Windows, Unix	Arquitectura cliente/servidor, pueden trabajar múltiples usuarios, posee control acceso y lenguaje de consultas
<b>Pulcinella</b>	Windows, Unix, Mac	Usa sistema de evaluación para calculo no probabilístico
<b>Sam Iam</b>	Windows, Unix, (Java ejecutable)	Solo hace análisis de sensibilidad
<b>WinMine</b>	Windows	Aprendizaje de la estructura



**Anexo 4.** Clasificación de Software de redes bayesianas y clasificadores bayesianos en software propietario**Software propietario**

**AgenaRisk**, herramienta visual que combina RB y simulación estadística, libre un mes para evaluación.

**Analytica**, basado en diagramas de influencia, ambiente visual para crear y analizar modelos probabilísticos. (Win/Mac).

**AT-Sigma Data Chopper**, para analizar y buscar relaciones causales en bases de datos.

**BayesiaLab**, herramienta de RB para aprendizaje supervisado y no supervisado, y una herramienta de análisis.

**Bayesware Discovery 1.0**, herramienta de modelación automática de RB desde datos buscando el modelo más probable.

**BNet**, incluye BNet.Builder para crear una RB, entrar información y obtener resultados y BNet.EngineKit para incorporar la tecnología RC (Redes de Creencia) a nuestras aplicaciones

**DXpress**, herramienta sobre Windows para crear y compilar RB.

**Ergo™**, Editor y resolvidor de RB (Win, Mac, demos disponibles).

**Flint**, combina RB, factores de certeza, y lógica difusa con un ambiente de programación lógica basado en reglas.

**Hugin**, colección completa de herramientas de razonamiento en RB.

**KnowledgeMiner**, usa redes neuronales autoorganizadas para descubrir la estructura del problema (Mac).

**Netica**, Herramienta de RB (Win 95-NT, demo disponible).

**PrecisionTree**, una macro de Microsoft Excel para crear árboles y diagramas de influencia.

**Anexo 5.** Clasificación de Software de redes bayesianas y Clasificadores Bayesianos en software libre

**Software libre**

**Bayda 1.0**, sistema experto para ecocardiografía.

**Bayesian belief network software**, de J. Cheng, incluye un *PowerConstructor*: Sistema eficiente para aprendizaje estructural y paramétrico de RB. Constantemente actualizado desde 1997 y un *PowerPredictor*: Programa de Minería de datos para modelación, clasificación y predicción de datos.

**Bayesian Logistic Regression Software**, regresión logística bayesiana a gran escala (Win y Linux).

**Bayesian Network tools in Java (BNJ)**, colección de código Fuentes de herramientas en java para aprendizaje y razonamiento probabilístico (Universidad del estado de Kansas, KDD Lab.).

**FDEP**, induce dependencia funcional desde una entrada de datos.

**GeNle**, ambiente de modelos de decisión mediante diagramas de influencia y RB (Win, tiene sobre 2000 usuarios).

**JavaBayes**, *software* de edición y uso de RB.

**jBNC**, conjunto de programas en Java para entrenamiento, prueba y aplicación de clasificadores de RB.

**JNCC, Naïve Credal Classifier 2**, herramienta en java que hace una extensión al Naïve bayes con resultados robustos aun cuando se tengan pequeños conjuntos de datos y/o información incompleta.

**MSBN: Microsoft Belief Network Tools**, herramienta para crear y evaluar RC bayesianas (libre para investigaciones no comerciales).

**PNL**, librería de código Fuentes de RB.

**Pulcinella**, herramienta para propagar incertidumbre basada en cálculos locales (*Lisp*).

**Anexo 6. Entrevista a profundidad guiada por las siguientes interrogantes:**

1. ¿Considera necesario que las instituciones agropecuarias en Cuba dispongan de base de datos con información confiable relativa a las enfermedades ganaderas que con frecuencia afectan la ganadería bovina?
2. ¿Considera importante la incorporación de las TIC para mejorar la efectividad de los diagnósticos que se realizan?
3. ¿Considera que los métodos utilizados para el procesamiento de la información emplean estrategias eficientes que producen información significativa?
4. ¿Se realizan análisis de las relaciones de dependencia entre las variables?
5. ¿Qué opinión le merece el SIVE?
6. Mencione otras mejoras que necesite realizar el proceso de diagnóstico de enfermedades en la ganadería y en particular el de la Fasciolosis bovina.
7. ¿Valora necesaria y pertinente el tema de investigación que está desarrollando?

**Anexo 7. Encuesta destinada a los especialistas de salud animal de los departamentos de las instituciones agropecuarias donde se gestione información sobre el comportamiento de las enfermedades ganaderas que con frecuencia afectan la ganadería en la provincia de Pinar del Río.**

Buenos días (tardes):

Actualmente se está desarrollando una investigación sobre el diagnóstico de enfermedades que con frecuencia afectan la masa ganadera, en particular la Fasciolosis bovina con el objetivo de brindar información con mayor grado de certeza para apoyar la toma de decisiones. Por tal motivo, se ha considerado indispensable recopilar información de las instituciones agropecuaria que incluyen procesos de análisis de estas enfermedades.

Usted ha sido seleccionado como posible experto, para ser consultado respecto a temas relacionados con la investigación que se está llevando a cabo. Agradecemos sinceramente su valiosa cooperación. Sus respuestas serán confidenciales.

**Muchas gracias por su contribución.**

**1. DATOS PERSONALES DEL ENCUESTADO**

Institución:	Profesión:
Años de experiencia en la profesión:	

**2. ELABORACIÓN DE DIAGNÓSTICO DE ENFERMEDADES**

**2.1.** ¿En su institución utilizan alguna herramienta para el procesamiento de la información y ejecución de diagnóstico de enfermedades en la ganadería?

Sí  No  No sé

¿Cuál(es)? \_\_\_\_\_

2.2. Marque con una X la frecuencia con que participa en la elaboración conjunta de diagnóstico de enfermedades en la ganadería en su institución.

No.	Frecuencia	Selección
1	Nunca	
2	Pocas veces	
3	A veces	
4	Muchas Veces	

2.3. Marque con una X la frecuencia con la que utiliza la información que se procesa relativa a las enfermedades ganaderas para la elaboración de diagnóstico.

No.	Frecuencia	Selección
1	Nunca	
2	Pocas veces	
3	A veces	
4	Muchas Veces	

2.4. ¿Su institución puede hacer uso de la información relativa a las enfermedades en la ganadería, de otras instituciones agropecuarias?

\_\_\_ Sí \_\_\_ No

### 3. ALMACENAMIENTO DE LA INFORMACIÓN

3.1. Marque con una X la frecuencia con que se requiere emplear información para realizar diagnóstico de enfermedades en la ganadería.

No.	Frecuencia	Selección
1	Nunca	
2	Pocas veces	
3	A veces	
4	Muchas Veces	

3.2. Marque con una X a través de que vía comparte las informaciones para realizar diagnósticos de enfermedades con otros especialistas de salud animal en su institución.

- Correo electrónico
- Dispositivo externo (memoria flash, CD, entre otros)
- Base de datos de información relacionada con las enfermedades ganaderas que con frecuencia afectan la ganadería bovina
- Ninguno
- Otros ¿Cuáles? \_\_\_\_\_

**3.3.** ¿Dónde se almacena la información relativa a las enfermedades ganaderas que con frecuencia afecta la ganadería, en su institución?

No.	Categoría	Selección
1	Base de datos para almacenar información sobre las enfermedades ganaderas que con mayor frecuencia afectan la ganadería	
2	Web personal de cada especialista de salud animal	
3	Un servidor central o intranet	
5	Un servidor FTP	
6	No se almacena	
7	No sé	

**3.4.** Considera necesario que su institución posea alguna herramienta que favorezca el almacenamiento y la gestión de los datos relativos a las enfermedades ganaderas.

Sí  No

#### **4. UTILIDAD DE LA INFORMACIÓN**

**4.1.** Considera favorable realizar intercambios de experiencias frecuentes sobre las enfermedades ganaderas que con frecuencia afectan la ganadería, entre los especialistas de salud animal de diversas instituciones.

**4.2.** ¿Considera que sería importante mejorar la gestión, almacenamiento y análisis de la información para el diagnóstico de enfermedades en la ganadería en Cuba para garantizar una mejor calidad en los diagnósticos?

Sí  No

**4.3.** ¿Considera que los métodos utilizados para el procesamiento de la información emplean estrategias eficientes que producen información significativa?

Sí  No

**4.4.** ¿Conoce alguna institución de salud animal cubana que haga uso de la información generada desde el SIVE, para ejecutar diagnóstico de enfermedades?

Sí  No

Si la respuesta es Sí, menciónelas. \_\_\_\_\_

**Anexo 8. Encuesta destinada a los especialistas en Medicina Veterinaria que poseen dominio informático en cuanto al procesamiento de la información relacionada con la vigilancia epizootiológica.**

Buenos días (tardes):

Actualmente se está desarrollando una investigación sobre el diagnóstico de enfermedades que con frecuencia afectan la masa ganadera, en particular la Fasciolosis bovina con el objetivo de brindar información con mayor grado de certeza para apoyar la toma de decisiones. Por tal motivo, se ha considerado indispensable recopilar información de las instituciones agropecuaria que incluyen procesos de análisis de estas enfermedades.

Usted ha sido seleccionado como posible experto para ser consultado respecto a temas relacionados con la investigación que se está llevando a cabo. Agradecemos sinceramente su valiosa cooperación.

Sus respuestas serán confidenciales.

**Muchas gracias por su contribución.**

**1. DATOS PERSONALES DEL ENCUESTADO**

Institución:	Profesión:
Años de experiencia en la profesión:	

2. ¿En su institución utilizan alguna herramienta para el almacenamiento y análisis de datos relativos a las enfermedades ganaderas que con frecuencia afectan la ganadería en Cuba?

Sí  No  No sé

¿Cuál(es)? \_\_\_\_\_



### 3. ALMACENAMIENTO Y DISTRIBUCIÓN DE LA INFORMACIÓN RELACIONADA CON LAS ENFERMEDADES GANADERAS QUE CON FRECUENCIA AFECTAN LA GANADERÍA CUBANA.

3.1. ¿Dónde se almacena la información, de su institución de referencia?

No.	Categoría	Selección
1	Base de datos para almacenar información sobre las enfermedades ganaderas que con mayor frecuencia afectan la ganadería bovina	
2	Web personal de cada especialista de salud animal	
3	Un servidor central o intranet	
5	Un servidor FTP	
6	No se almacena	
7	No sé	

3.2. ¿Tiene previsto su institución crear una base de datos para almacenar la información relacionada con las enfermedades ganaderas que con frecuencia afectan la ganadería cubana?

Sí  No  No tengo conocimiento

3.3. Cree usted necesario que su institución haga uso de alguna herramienta que permita almacenar y analizar la información relativa a las enfermedades ganaderas.

Sí  No

3.4. ¿Conoce alguna otra institución de salud animal, cubana, que haga uso de bases de datos para almacenar la información relativa a las enfermedades ganaderas?

Sí  No

Si la respuesta es Sí, menciónala(s). \_\_\_\_\_

3.5. Marque con una X la(s) herramienta(s) utilizada(s) en su institución para almacenar los datos de las enfermedades ganaderas que con frecuencia afectan la ganadería en Cuba.

SIVE

Desarrollo propio. ¿Cuál? \_\_\_\_\_

Ninguno

Otros \_\_\_\_\_

#### 4. PROCESAMIENTO DE LA INFORMACIÓN

4.1. Marque con una X cuáles son los problemas en los datos, que con frecuencia se presentan:

- Datos distorsionados
- Datos que no fueron incluidos
- Datos irrelevantes
- Diferentes tipos de datos
- Características faltantes o insuficientes
- Mucha cantidad de datos
- Una sola fuente de datos
- Ninguno
- No sé
- Otros. ¿Cuáles? \_\_\_\_\_

#### Anexo 9. Entrevistas en profundidad para valorar los beneficios y aportes del modelo MRB-DiagPron.

1. Experiencia de los entrevistados en los procesos de diagnóstico de enfermedades en la ganadería.
2. Necesidades actuales de mejorar el diagnóstico de enfermedades en la ganadería.
3. Conocimiento de modelos, métodos, procedimientos o técnicas utilizadas hoy para el diagnóstico de enfermedades en la ganadería.
4. Conocimiento de la aplicación de elementos en el entorno nacional o internacional para mejorar diagnóstico de enfermedades en la ganadería.
5. Pertinencia y novedad que se le confiere al modelo desarrollado, a partir de los principios establecidos y componentes desarrollados.
6. Beneficios de la aplicación del modelo en el Sistema Nacional Veterinaria.

7. Valoración de la manera en que son analizadas las habilidades técnicas y no técnicas del personal de vigilancia epizootiológica de las instituciones agropecuarias, en el diagnóstico de enfermedades en la ganadería.
8. Valoración de la utilidad de la información que brinda el modelo para facilitar la toma de decisiones de los especialistas de salud animal, a través de su instanciación en la herramienta informática desarrollada.