CURSO 2018-2019

ENTREGABLE REDES BAYESIANAS

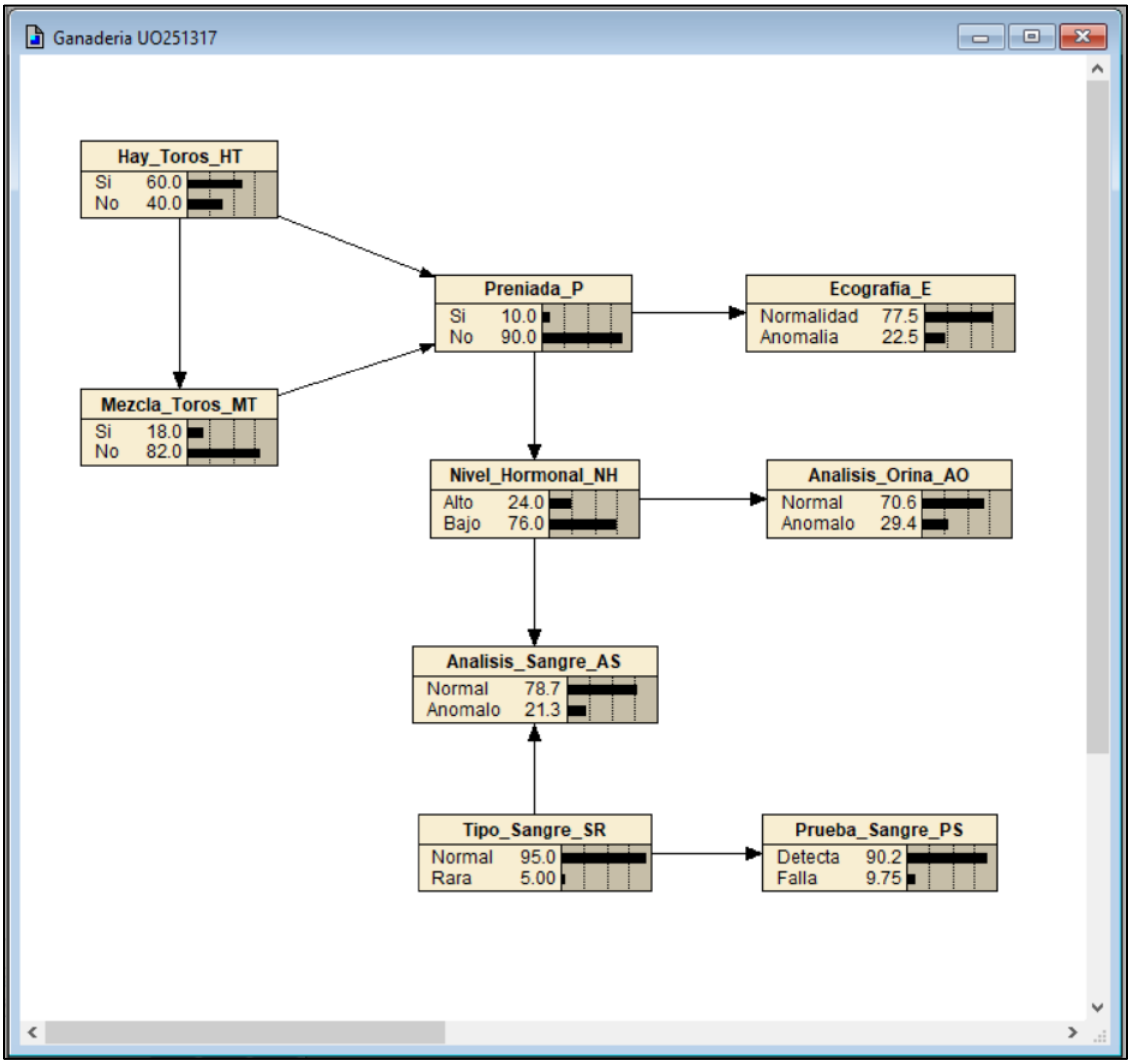
Autor:

- José Antonio García García - UO251317

Contenido

1. RED BAYESIANA

Esta es la red bayesiana que he deducido a partir de las especificaciones del enunciado.

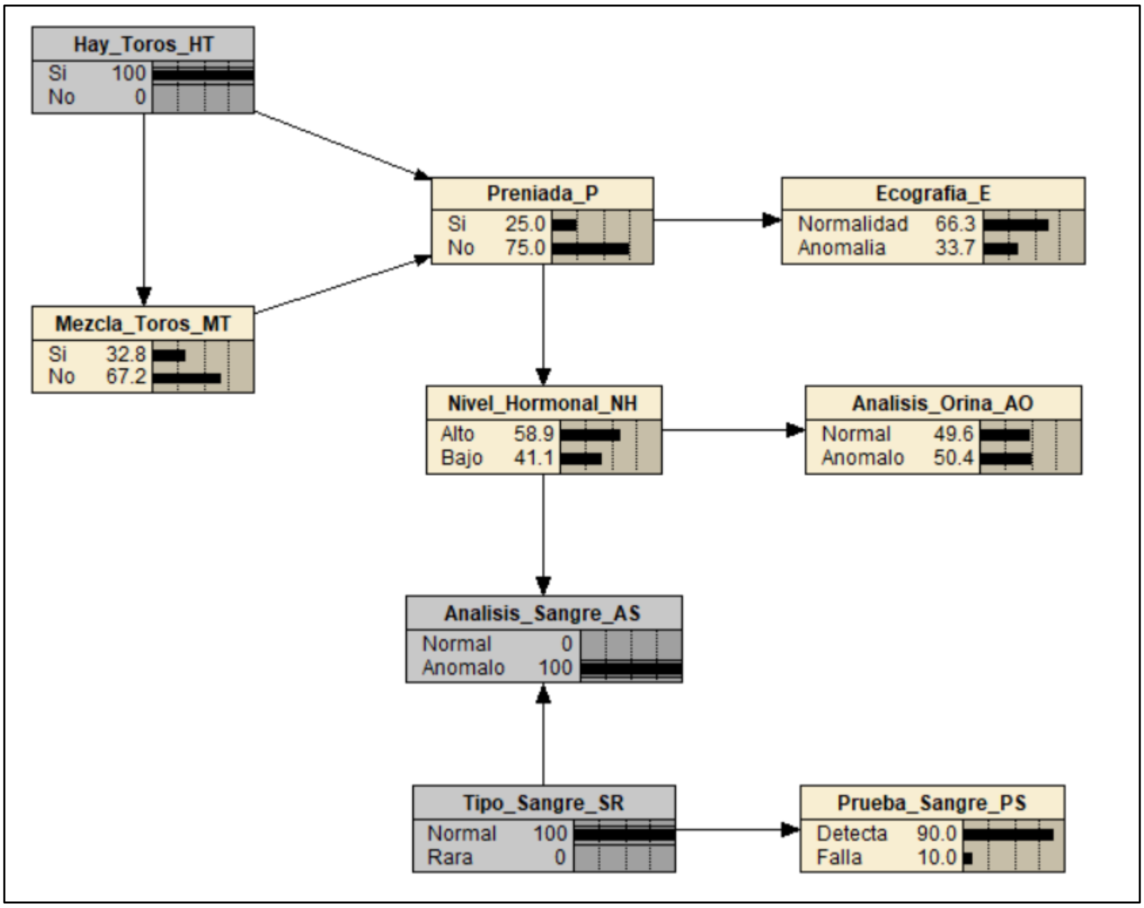


2. EJERCICIOS

EJERCICIO 3: SUBAPARTADO 1.

Si el granjero tiene toros, la vaca no tiene sangre "rara" y su muestra de sangre es anómala, ¿cuál es la probabilidad de que la vaca esté preñada?

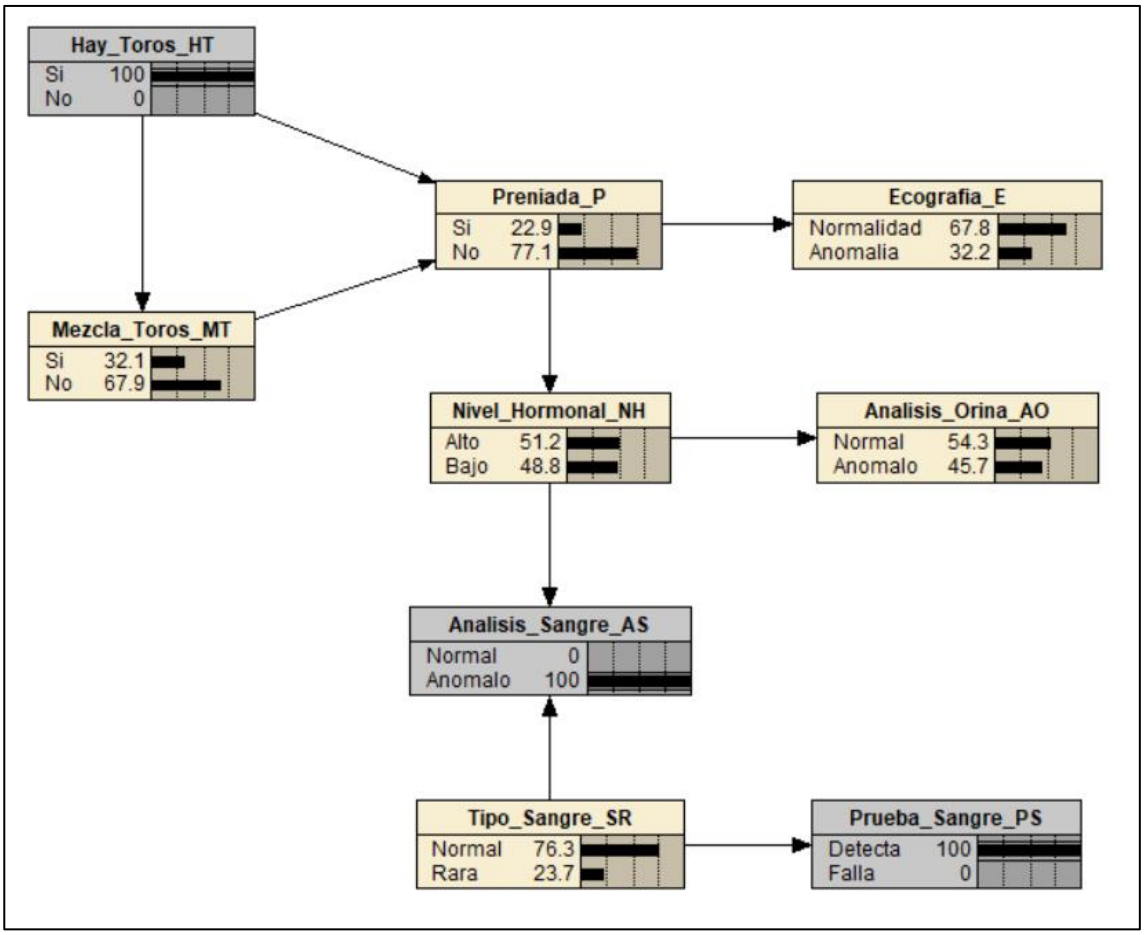
La probabilidad de que la vaca estuviera preñada en esta situación es del 25%.



EJERCICIO 3: SUBAPARTADO 2.

¿Y si no sabemos nada acerca de la sangre rara, pero el test dice que sí la tiene?

Si se diera esta situación, la probabilidad de que la vaca estuviera preñada es del 22.9%



EJERCICIO 4.

Apoyándote en la ventana de mensajes de Netica, crea la distribución conjunta para el subconjunto de variables (MT, E, AS).

La siguiente tabla contiene todas las posibles combinaciones de estas tres variables. Para la realización de este ejercicio consideraremos verdadero MT si hay toros, a E si la ecografía fue con normalidad y a AS si es resultado es normal.

MT	E	AS	Resultado
T	T	T	8.916
T	T	F	2.334
T	F	T	4.832
T	F	F	1.918
F	T	T	52.815
F	T	F	13.435
F	F	T	12.097
F	F	F	3.653
			100.000

EJERCICIO 5: SUBAPARTADO 1.

Elige un valor para cada una de las variables y calcula a mano la probabilidad (transparencia 42 de teoría): $P\ (HT, \neg MT, P, \neg E, \neg NH, \neg AO, AS, \neg SR, \neg PS)$.

$P = (HT, \sim MT, P, \sim E, \sim NH, \sim AO, AS, \sim SR, \sim PS) =$

$= P(HT) \cdot P(\sim MT|HT) \cdot P(P|\sim MT, HT) \cdot P(\sim E|P) \cdot P(\sim NH|P)$

$\cdot P(\sim AO|\sim NH) \cdot P(AS|\sim NH, \sim SR) \cdot P(\sim SR) \cdot P(\sim PS|SR) =$

$= 0{,}6 \cdot 0{,}7 \cdot 0{,}1 \cdot 0{,}9 \cdot 0{,}6 \cdot 0{,}75 \cdot 0{,}57 \cdot 1{,}00 \cdot 0{,}1 =$

$= 0{,}00096957 \underset{x100\%}{\Rightarrow} 0{,}096957\ \%$

EJERCICIO 5: SUBAPARTADO 2.

Explica cómo sería el procedimiento si lo que quisieses calcular fuese la probabilidad condicionada (transparencia 43 de teoría): $P\ (AS|\ AO, \neg P)$.

Para el cálculo de la probabilidad condicionada tendríamos que hacer el mismo procedimiento que en el apartado anterior: calcular la probabilidad de que se diera el escenario que queremos probar. En este caso, al no tener una combinación con el estado de todas las variables, sería necesario comprobar la probabilidad de todas las soluciones, es decir: $\boldsymbol{AS}$, $\boldsymbol{AO}$, $\boldsymbol{\neg P}$ como predefinidas y luego realizar una tabla como la del ejercicio 4 con toda la combinatoria, y una segunda iteración con $\boldsymbol{\neg AS}$, $\boldsymbol{AO}$, $\boldsymbol{\neg P}$. Esto produciría 128 combinaciones distintas de las variables del problema.

EJERCICIO 6: SUBAPARTADO 1.

Siguiendo la condición de Markov, ¿qué valores debería conocer si quiero afirmar que E y PS son independientes? ¿Y para afirmar que MT y AS lo son?

En el caso de **E** y **PS** se considerarían independientes si conociéramos a sus respectivos padres: para que **E** sea independiente necesitamos conocer el valor de **P** mientras que para que **PS** sea independiente necesitamos conocer el valor de **SR**.

Por otro lado, en el caso de **MT** y **AS**, necesitaríamos conocer los valores de **HT** para **MT** y los valores de **NH** y **SR** para el caso de **AS**.

EJERCICIO 6: SUBAPARTADO 2.

Apoyándote en el concepto de D-separación, ¿crees que P y AS son independientes? Explica por qué.

P y **AS** no son independientes ya que entre ellos existe el camino **P -> NH -> AS**, formado por flechas de tipo no cabeza-cabeza, y **NH** no está observado.

EJERCICIO 6: SUBAPARTADO 3.

Apoyándote en el concepto de D-separación, ¿crees que P y PS son independientes? Explica por qué. ¿Y si conociésemos AS?

En este caso **P** y **PS** si serian independientes ya que existe un bloqueo en las observaciones del nodo **AS** de tipo cabeza-cabeza no siendo este observado.

Esto cambiaria si observamos **AS**, porque dejaría de ser un bloqueo, lo que crearía dependencia entre los nodos **P** y **PS**.

EJERCICIO 7.

Dar una traza de ejemplo de cómo generar una posible muestra estocástica para calcular $P\,(MT|\,E,\neg AS)$. Deberás indicar, en cada paso, el número aleatorio que ha salido y por qué has asignado cada valor a cada variable.

Para empezar, es necesario tener los nodos en orden topológico, es decir, ordenar al principio los que no tengan padres y luego todos los nodos de los que tengamos ambos padres ya ordenados. De esta forma nos quedaría la siguiente lista.

HT	**SR**	**MT**	**PS**	**P**	**E**	**NH**	**AO**	**AS**

A continuación, vamos a obtener los valores aleatorios que necesitamos para realizar la traza. Son los siguientes.

0,42	**0,18**	**0,61**	**0,97**	**0,06**	**0,24**	**0,85**	**0,59**	**0,91**

Por último, escogeremos los valores que utilizaremos en cada caso para calcular el resultado.

HT: Si cuando el número sea inferior al porcentaje de Si.

SR: Normal cuando el número sea inferior al porcentaje de Normal.

MT: Si cuando el número sea inferior al porcentaje de Si.

PS: Detecta cuando el número sea inferior al porcentaje de Detecta.

P: Si cuando el número sea inferior al porcentaje de Si.

E: Normalidad cuando el número sea inferior al porcentaje de Normalidad.

NH: Alto cuando el número sea inferior al porcentaje de Alto.

AO: Normal cuando el número sea inferior al porcentaje de Normal.

AS: Normal cuando el número sea inferior al porcentaje de Normal.

Con toda esta información la traza quedaría de la siguiente manera.

Nodo	Porcentaje	Aleatorio	Resultado
HT	0,60	0,42	Si
SR	0,95	0,18	Normal
MT	0,30	0,61	No
PS	0,90	0,97	Falla

P	0,10	0,06	Si
E	0,10	0,24	Anomalía
NH	0,60	0,85	Bajo
AO	0,85	0,59	Normal
AS	0,90	0,91	Anómalo

EJERCICIO 8.

Repetir el apartado anterior, pero utilizando el método de ponderación de la verosimilitud. Se deberá utilizar la misma secuencia de números aleatorios que en el caso anterior. $P\ (MT|\ E, \neg AS)$.

Primero, vamos a volver a apuntar los datos que tenemos:
- La lista de nodos ordenados.
- La lista de números aleatorios que vamos a utilizar.

HT	SR	MT	PS	P	E	NH	AO	AS
0,42	0,18	0,61	0,97	0,06	0,24	0,85	0,59	0,91

Empezando con W = 1, la traza seria de la siguiente manera.

Nodo	Porcentaje	Aleatorio	Resultado	W
HT	0,60	0,42	Si	1
SR	0,95	0,18	Normal	1
MT	0,30	0,61	No	1
PS	0,90	0,97	Falla	1
P	0,10	0,06	Si	1
E	0,10	*Forzado*	Normalidad	0.1
NH	0,60	0,85	Bajo	0.1
AO	0,85	0,59	Normal	0.1
AS	0,90	0,91	Anómalo	0.1

Este seria el resultado de la traza con un peso de W = 0,04. En el caso de que quisiéramos sacar la probabilidad tendríamos que repetir esta secuencia de pasos N veces con valores aleatorios diferentes cada vez.

EJERCICIO 9.

Por último, indica qué ventajas tiene el método de ponderación de la verosimilitud sobre el muestreo estocástico.

La principal ventaja que tiene el método de ponderación de la verosimilitud frente al muestreo estocástico es que, dado un nodo del cual queremos calcular la probabilidad en un caso concreto, si la probabilidad de que dicha situación ocurra es muy baja o casi cero, el número de iteraciones que tendríamos que realizar seria enorme, siendo muchas de estas despreciables por no cumplir el caso que queremos probar. Es por esta razón que el método de forzar los valores de la verosimilitud permite que, aun teniendo probabilidades muy cercanas a 0, se puedan obtener valores relevantes para el cálculo de probabilidades.