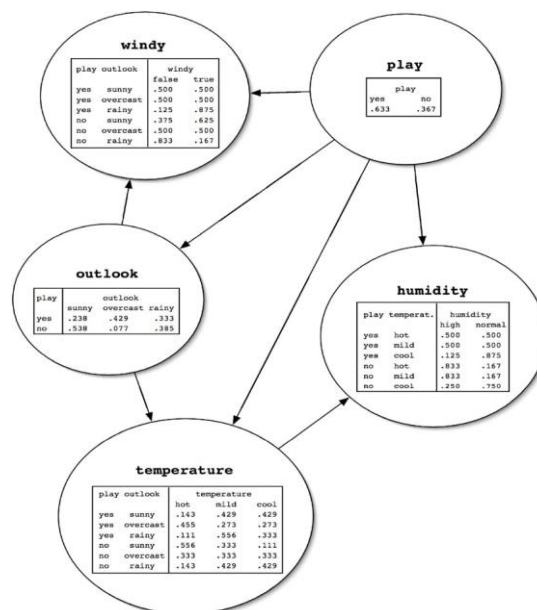


Práctica 7.2: Ampliación Métodos Bayesianos

Alejandro Pulido Sánchez

Ejercicios y prácticas Métodos Bayesianos II, aprendizaje de parámetros.
Curso 2022-2023

1. Considerar el conjunto de datos weather.nominal.practica proporcionado y la siguiente red TAN:



Obtener los parámetros de la red para el nodo Windy utilizando la estimación de máxima verosimilitud y la corrección de Laplace con $m = 0,5$ (caso particular de la estimación Bayesiana). Hacer el ejercicio sin utilizar software de aprendizaje.

$$\hat{\theta}_{x_i^j|u} = \frac{M[u, x_i^j] + mp}{M[u] + m}$$

(Sunny, yes) → Total de instancias: 2 ($M[u]$)

- » TRUE → $(1 + 0.5 * 0.5) / (2 + 0.5) = 0.5$
- » FALSE → $(1 + 0.5 * 0.5) / (2 + 0.5) = 0.5$

(Sunny, no) → Total de instancias: 3 ($M[u]$)

- » Para TRUE → $(1 + 0.5 * 0.5) / (3 + 0.5) = 0.3571$
- » Para FALSE → $(2 + 0.5 * 0.5) / (3 + 0.5) = 0.6428$

(Overcast, yes) → Total de instancias: 2 (M[u])

- » Para TRUE → $(2 + 0.5 * 0.5) / (2 + 0.5) = 0.9$
- » Para FALSE → $(0 + 0.5 * 0.5) / (2 + 0.5) = 0.1$

(Overcast, no) → Total de instancias: 0 (M[u])

- » Para TRUE → $(0 + 0.5 * 0.5) / (0 + 0.5) = 0.5$
- » Para FALSE → $(0 + 0.5 * 0.5) / (0 + 0.5) = 0.5$

(Rainy, yes) → Total de instancias: 5 (M[u])

- » Para TRUE → $(1 + 0.5 * 0.5) / (5 + 0.5) = 0.2272$
- » Para FALSE → $(4 + 0.5 * 0.5) / (5 + 0.5) = 0.7727$

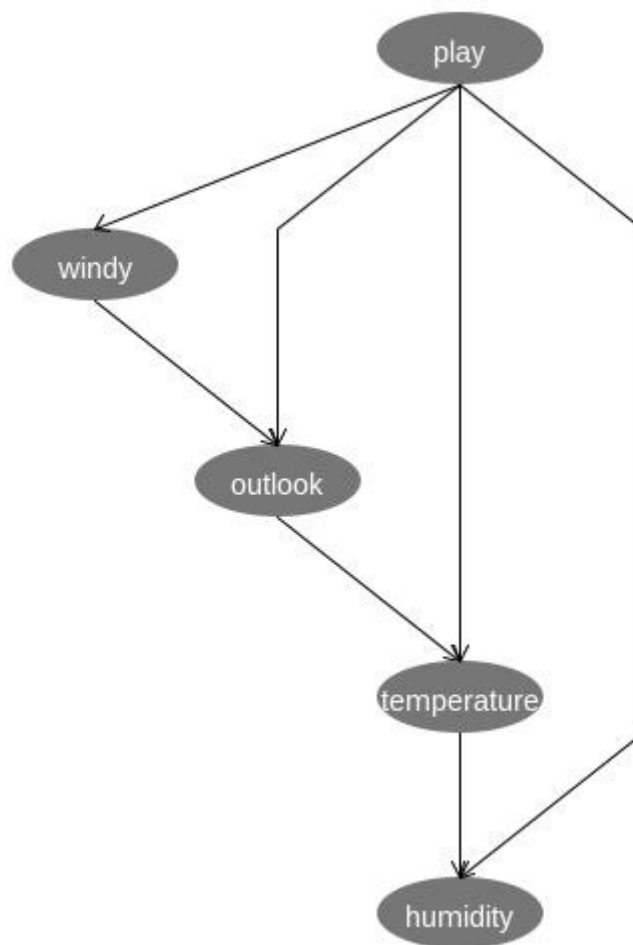
(Rainy, no) → Total de instancias: 2 (M[u])

- » Para TRUE → $(2 + 0.5 * 0.5) / (2 + 0.5) = 0.9$
- » Para FALSE → $(0 + 0.5 * 0.5) / (2 + 0.5) = 0.1$

	(sunny,yes)	(sunny,no)	(overcast,yes)	(overcast,no)	(rainy,yes)	(rainy,no)
TRUE	0,5	0,3571	0,9	0,5	0,2272	0,9
FALSE	0,5	0,6428	0,1	0,5	0,7727	0,1

2. Considerar el conjunto de datos weather.nominal.practica proporcionado y Entrenar con Weka una red TAN. Comparar ambas redes. Si la estructura es la misma que la proporcionada en el ejercicio anterior, comparar también los parámetros de la red para el nodo Windy.

Tras obtener la red TAN en weka, se puede observar que el algoritmo no toma en cuenta la dirección del grafo, por lo que la estructura es la misma.

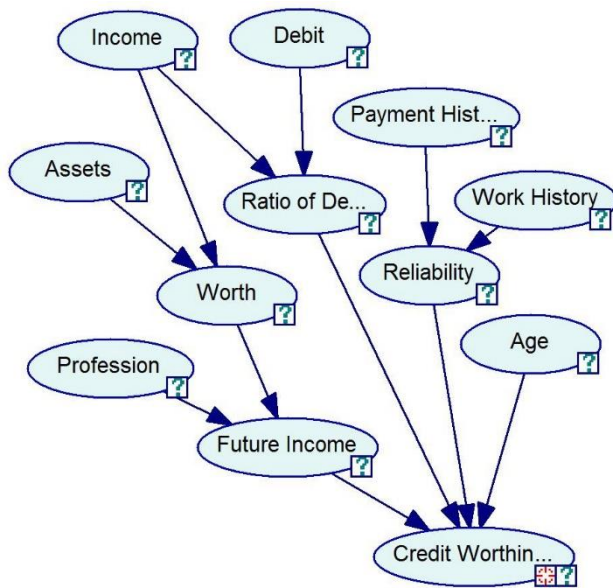


Vemos que la flecha de windy a outlook en el enunciado se encuentra en el sentido contrario. Esto hace que los parámetros del nodo windy y outlook sean distintos a los dados, ya que ahora tienen nodos padres distintos. Obteniendo, ahora, los siguientes resultados:

Probability Distribution Table For windy ×		
play	TRUE	FALSE
yes	0,45	0,55
no	0,583	0,417

Probability Distribution Table For outlook ×			
play windy	sunny	overcast	rainy
yes TRUE	0,273	0,455	0,273
yes FALSE	0,231	0,077	0,692
no TRUE	0,333	0,111	0,556
no FALSE	0,714	0,143	0,143

3. Considerar la estructura la red Bayesiana Credit, en la que todas las variables son discretas:



A simple network for assessing credit worthiness of an individual, developed by Gerardina Hernandez as a class homework at the University of Pittsburgh. Note that all parentless nodes are described by uniform distributions. This is a weakness of the model, although it is offset by the fact that all these nodes will usually be observed and the network will compute the probability distribution over credit worthiness correctly. Another element of this model is that only the node CreditWorthiness is of interest to the user and is designated as a target.

Ejemplo Credit, proporcionado por la herramienta GeNIe2.0, cuya versión de educación ya no está disponible.

Se proporcionan tres conjuntos de datos, Datos_Credit_100.csv, Test_Credit_1000.csv y Datos_Credit_10000.csv, que contienen 100, 1000, y 10000 instancias, respectivamente, todas ellas muestreadas de forma aleatoria e independiente a partir de la red. Por tanto, todos los ejemplos son muestras independientes de la distribución de probabilidad modelada por la red, P^{credit} .

Se desea examinar el comportamiento de las redes Naive Bayes y TAN sobre los conjunto de datos Datos_Credit_100.csv y Datos_Credit_10000.csv.

Utilizar Weka para entrenar los dos tipos de redes con los dos conjuntos de datos. En todos los casos, estimar los parámetros con la estimación de máxima verosimilitud con corrección de Laplace ($\alpha = 0,5$) y estimar la tasa de error de cada modelo mediante validación cruzada de diez particiones.

Una vez generado cada modelo y estimada su tasa de error mediante validación cruzada sobre los conjuntos proporcionados, guardar el modelo y estimar su tasa de error sobre el conjunto Test_credit_1000, que contiene otras mil instancias muestreadas aleatoriamente de la red.

Comparar el comportamiento de las redes en base a la estimación de la tasa de error sobre Test_Credit_1000. Comparar esta estimación con la que proporciona la estimación mediante validación cruzada.

	Datos_Credit_100		Datos_Credit_10000	
	Naive Bayes	Red TAN	Naive Bayes	Red TAN
Cross Validation 10	36%	35%	29,71%	28,27%
Test_Credit_1000	30%	33,20%	27,50%	25%

En la tabla podemos observar las tasas de error obtenidas de realizar los dos algoritmos (NB y Red TAN) con dos métodos de entrenamiento distintos. Uno se ha hecho mediante validación cruzada de 10 particiones y otro usando Test_credit_1000 como fichero de entrenamiento.

Se observa que al aumentar el tamaño del conjunto de entrenamiento disminuye ligeramente la tasa de error. Si comparamos los resultados de los dos algoritmos vemos que, a priori, para una cantidad mayor en el numero de datos de entrenamiento, el algoritmo de red TAN funciona mejor que Naive Bayes. Pero teniendo en cuenta que las variaciones son pequeñas y la cantidad de resultados experimentales es ínfima, no podemos sacar ninguna conclusión certera.

ANEXO: aprendizaje de redes bayesianas en Weka

En Weka, los algoritmos de aprendizaje de estructura de la red se encuentran en `weka.classifiers.bayes.Bayesnet`. Aparece un cuadro de dialogo con varias opciones. Nos interesan las opciones:

Estimator: método de estimación de parámetros. Por defecto “*SimpleEstimator –A 0.5*” que se corresponde con la estimación de máxima verosimilitud con corrección tipo Laplace y $m=0.5$ (caso particular de la estimación Bayesiana).

searchAlgorithm:

- Pinchando sobre *Choose* se selecciona el algoritmo de búsqueda de la estructura:
 - Naive Bayes: seleccionar *fixed.NaiveBayes*
 - Tan: seleccionar *local.TAN*
- Pinchando sobre los valores de *Choose* se configura cada algoritmo:
 - NaiveBayes: no hay opciones adicionales.
 - TAN: dejar los parámetros por defecto. La puntuación Bayes con *alfa=0,5* es la estimación de máxima verosimilitud con la corrección tipo Laplace.