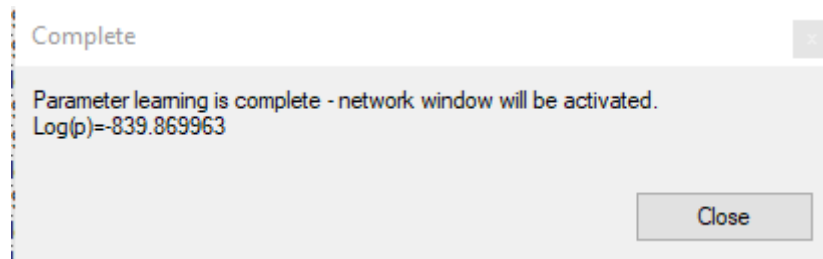


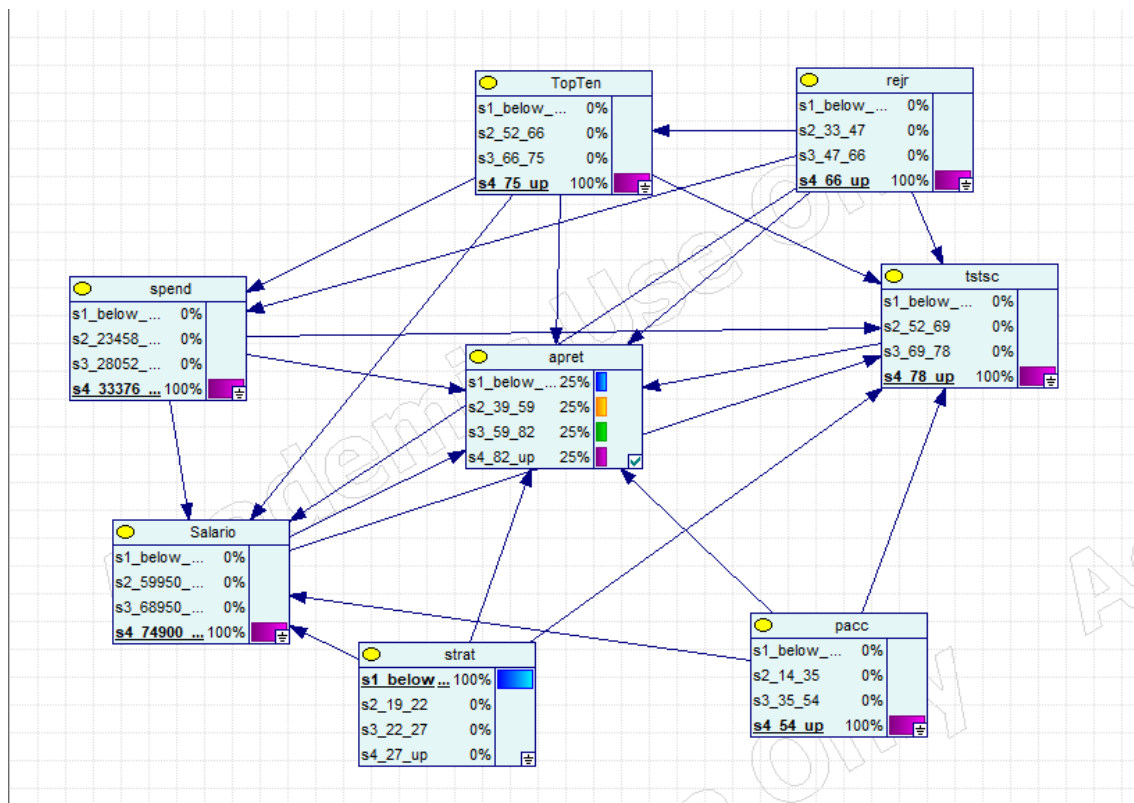
Práctica 6 – Aprendizaje con GeNIe

Pregunta 1: Indica el valor de $\log(p)$ que te ha dado en el proceso de aprendizaje.



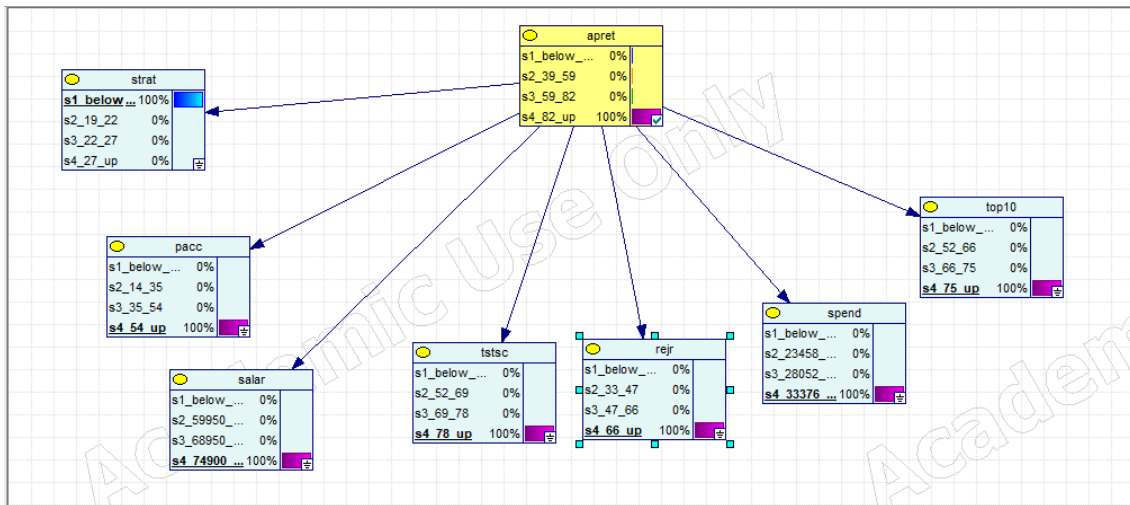
El valor de $\log(p)$ = -839,869963...

Pregunta 2: Supón un nuevo ejemplo en el que Tstsc toma el valor de más de 78; Top 10 más de 75); Pacc más de 54, Spend más de 33376, Strat menos de 19; Salar más de 74900 y Rejr más de 66. Calcula la probabilidad del nodo Apret para este ejemplo (captura la pantalla que muestra las probabilidades), y di cómo se clasificaría este ejemplo.



Se establecen las probabilidades que se observan en la imagen, es decir, para cada rango de valores que puede tomar Apret hay una probabilidad de 1/4 de que lo tome. Con este resultado no podemos clasificar este ejemplo, pues son sucesos equiprobables.

Pregunta 3: Repite lo pedido en la pregunta 2, con el modelo Naive Bayes obtenido.



Se establecen estas probabilidades para el nodo Apret y en este caso nos sale una probabilidad de uno para el rango de valores "mayor que 82". Por tanto podemos clasificar este ejemplo diciendo que Apret tomará un valor mayor a 82 con los datos que se nos dan.

Pregunta 4: Escribe un breve informe acerca de la calidad del modelo aprendido, tanto para el caso de redes bayesianas como el modelo Naive Bayes (incluye también los valores obtenidos para el área bajo la curva ROC, en ambos casos). A la vista de los resultados obtenidos, ¿qué modelo es mejor?

En el caso de modelo con redes bayesianas se ha obtenido un porcentaje de exactitud (accuracy) del 83.77%, lo cual indica un gran nivel de acierto. Tenemos que en cuanto a las hobexas, el 65.7% están bien clasificadas, 895 hobexas clasificadas por el modelo son realmente hobexas y con la ayuda de la matriz de confusión 466 son wackas. En cuanto a las wackas tenemos un porcentaje de acierto de 86.18%, 1142 wackas son realmente wackas, 130 son hobexas y 53 son wurras. Para el caso de los wurras, todos los wurras clasificados como tal lo son realmente. El área debajo de la curva que obtenemos es $AUC=0.939454$.

En el caso de modelo con redes bayesianas se ha obtenido un porcentaje de exactitud (accuracy) del 82.27%, lo cual indica un gran nivel de acierto también. Tenemos que en cuanto a las hobexas, el 79.79% están bien clasificadas, 1086 hobexas clasificadas por el modelo son realmente hobexas y con la ayuda de la matriz de confusión 275 son wackas. En cuanto a las wackas tenemos un porcentaje de acierto de 67.24%, 891 wackas son realmente wackas, 381 son hobexas y 53 son wurras. Para el caso de los wurras, todos los wurras clasificados como tal lo son realmente. El área debajo de la curva que obtenemos es $AUC=0.939117$.

Con todo esto, vemos que son modelos que prácticamente dan resultados muy parecidos. Por tanto, si debemos elegir cual puede ser el mejor, diríamos que es el que usa el modelo de redes bayesianas, ya que tiene un AUC mayor, lo que significa que la curva de ROC tiene una mejor posición. Sin embargo son áreas muy próximas.