PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA
DEPARTAMENTO DE CIENCIA DE LA COMPUTACIÓN
IIC2613 INTELIGENCIA ARTIFICIAL

Noviembre 2014

Interrogación 3

Tiempo: 90 minutos, SIN APUNTES

Nombre:_______________________________

1. (*20 puntos*) **Responda las siguientes preguntas fundamentando brevemente sus respuestas.**

a. En el contexto de MDPs, una condición para que el algoritmo de value iteration encuentre la política óptima es que para todo estado la probabilidad de transitar a cualquiera de los estados siguientes no dependa de los estados anteriores.

   Sol: Verdadero, el supuesto de markovianidad implica la no dependencia en estados anteriores, por tanto, es un requisito para que los algoritmos converjan a una política óptima.

b. En una red neuronal tipo feed-forward entrenada con el algoritmo de backpropagation, al aumentar significativamente el número de ejemplos del set de entrenamiento generalmente el error en el set de validación disminuye y en el set de entrenamiento aumenta.

   Sol: Verdadero, generalmente un mayor set de entrenamiento implica mayores dificultades de modelación, por tanto, es esperable que el error de clasificación aumente o se mantenga. Respecto al set de validación, dado que el entrenamiento se realiza con un set de datos mayor, i.e., más representativo al problema, es esperable obtener un modelo con mejores capacidades de generalización y por tanto que el error en este set disminuya.

c. En aprendizaje de máquina el espacio de hipótesis se refiere al conjunto de observaciones que se puede obtener en un cierto problema.

   Sol: Falso, el espacio de hipótesis se relaciona con el tipo de modelos usado para modelar las observaciones, no con el conjunto de observaciones.

d. El problema de sobreajuste en más factible en casos en que se cuenta con un espacio de hipótesis menor.

   Sol: Falso, mientras más limitadas sean las capacidades de modelación del espacio de hipótesis, más difícil será sobreajustar los datos, por tanto, un espacio de hipótesis menor tendrá menor probabilidad de sobreajustar.

e. El problema de sobreajuste es más factible cuando se cuenta con un menor set de datos de entrenamiento.

   Sol: Verdadero, generalmente un menor número de datos de entrenamiento implica menor representatividad y mayor facilidad de sobreajuste. Uno puede pensar en el caso extremo de sólo 1 dato de entrenamiento, lo cual es claramente fácil de sobreajustar.

f. El problema de sobreajuste es más factible cuando se cuenta con un menor set de datos de validación.

   Sol: Verdadero, el set de validación se utiliza para evitar el sobreajuste, por ejemplo, para podar un árbol de decisión o determinar un número adecuado de épocas para entrenar una red neuronal, por tanto, un set de validación empobrecido implica menor información para controlar el sobreajuste.

g. El problema de sobreajuste es más factible cuando se cuenta con un menor set de datos de test.

   Sol: Falso, una vez terminado el entrenamiento, el set de test se utiliza para obtener una estimación de las capacidades de generalización del modelo resultante, por tanto, no es usado para controlar sobreajuste, y por tanto, éste no depende de su tamaño.

h. En un problema de aprendizaje reforzado basado en el algoritmo de Q-learning la función de transición no puede ser modelada por una función Gausiana.

   Sol: Verdadero, Q-learning es una estrategia de aprendizaje de políticas de acción sin modelo.

i. En un MDP la probabilidad de llegar en el tiempo t+1 al estado S depende de la función de reward recibida en el estado t.

   Sol: Falso, hay un problema de edición en el enunciado deberia decir "tiempo t", de lo contrario el enunciado tiene poco sentido. Considerando esta corrección, el enunciado de todas formas es falso, pues la función de probabilidad de transición de estados depende sólo del estado actual y la acción elegida, $P(S'|S, a)$, no del valor de la recompenza.

j. En un problema de clasificación se cuenta con muchos datos representativos donde se probó el rendimiento de 3 algoritmos usando validación cruzada con 10-folds. Los resultados indican:

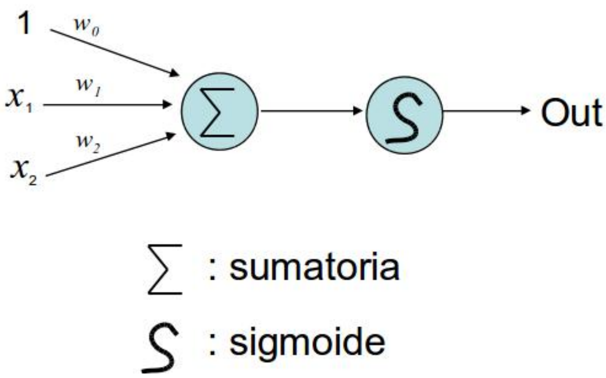
	Set de entrenamiento	Set de test
Algoritmo 1	94 %	92 %
Algoritmo 2	92 %	91 %
Algoritmo 3	90 %	93 %

   ¿Cuál algoritmo recomendaría?.

   Sol: recomendaría el algoritmo con mejor capacidad de generalización, i.e., mejor rendimiento en el set de test, algoritmo 3 .

2. (*16 puntos*) **Redes Neuronales**

a. (*12 pts*) En un problema de clasificación se decide utilizar la siguiente red neuronal:



Para el entrenamiento se inicializan los pesos de la red con los siguientes valores: $w_0 = 2$, $w_1 = 1$, $w_2 = -2$. Se decide utilizar como función de pérdida el error medio cuadrático, sin embargo, durante el entrenamiento de la red, esta función se aproxima mediante la estrategia incremental para actualización de los pesos de la red (gradiente estocástico). Si el primer dato de entrenamiento de la red es x_1 =0.6, x_2=0.8, y $out = 0$, ¿Cuál es el valor de los pesos después de usar su primera actualización usando este registro?. Considere que se utiliza un valor de 0.9 para la tasa de aprendizaje.
La función sigmoidal utilizada está dada por: $sigmoid(x) = \frac{1}{1+e^{-x}}$

Sol:
Para la entrada indicada y pesos iniciales, la salida de la red es:
$\sum_i x_i w_i = 2 * 1 + 1 * 0{,}6 - 2 * 0{,}8 = 1$
$sigmoid(\sum_i x_i w_i) = 0{,}7311$

La gradiente de aprendizaje es :
$\Delta\vec{w} = -0{,}9 \nabla E$

$$\nabla E = -(t_d - o_d)o_d(1 - o_d)\vec{x} = 0{,}1437 \begin{pmatrix} 1 \\ 0{,}6 \\ 0{,}8 \end{pmatrix}$$

$$\vec{w} \leftarrow \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} - 0{,}9 * 0{,}1437 * \begin{pmatrix} 1 \\ 0{,}6 \\ 0{,}8 \end{pmatrix} = \begin{pmatrix} 1{,}8706 \\ 0{,}9224 \\ -2{,}1035 \end{pmatrix}$$

b. (*4 pts*) En un problema de clasificación de imágenes bidimensionales de 256x256 pixeles, se utiliza una red neuronal profunda del tipo convolucional. La capa inicial de esta red consiste en la aplicación de 10 filtros de 7x7, los cuales son aplicados con un step horizontal de 3 pixeles y un step vertical de 5 pixeles. ¿Cuál es el número de parámetros a estimar en esta primera capa de la red?.

Sol:
Los parámetros corresponden a los pesos de cada filtro. Cada filtro tiene 7x7=49 pesos. Además hay un total de 10 filtror. Por tanto el número de parámetros de la primera capa de esta red es : 7x7x10=490.

3. (*16 puntos*) **Naive Bayes**

a. (*8 pts*) Suponga un problema de clasificación en que la clase objetivo puede tomar r_o posibles valores y existe un total de n variables binarias (atributos) disponibles para predecir el valor de esta clase. Asumiendo que se utilizan funciones multinomiales o binomiales para modelar todas las funciones de probabilidad relevantes, ¿Cuál es el número de parámetros que es necesario estimar en los siguientes casos?

- Estimación de $P(C|a_1, \ldots, a_n)$ usando la ecuación original de Bayes, i.e., resolver: $P(C|a_1, \ldots, a_n) \propto P(C) * P(a_1, \ldots, a_n|C)$

  Sol:
  Estimación $P(C)$ implica estimar $r_o - 1$ parámetros
  Estimación $P(a_1, \ldots, a_n|C)$ implica estimar $r_o * (2^n - 1)$ parámetros
  TOTAL: $r_0 * 2^n - 1$ parámetros.

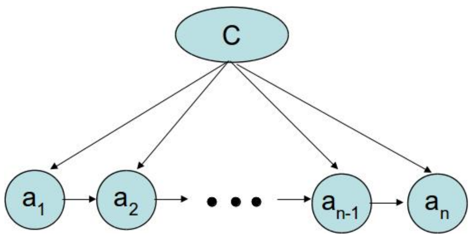
- Estimación de $P(C|a_1, \ldots, a_n)$ usando el supuesto de Naive Bayes, i.e., resolver: $P(C|a_1, \ldots, a_n) \propto P(C) * \prod_{i=1}^{n} P(a_i|C)$

  Sol:
  Estimación $P(C)$ implica estimar $r_o - 1$ parámetros
  Estimación $P(a_1, \ldots, a_n|C)$ implica estimar $r_o * n$ parámetros
  TOTAL: $r_o * (n + 1) - 1$ parámetros.

- En ambos casos anteriores, ¿Cuál es el número de parámetros a estimar cuando $r_o = 10$ y $n = 50$?

  Sol:
  Sin supuesto naive Bayes= $r_0 * 2^n - 1 = 10 * 2^{50} - 1 \approx 11 * 10^{15}$
  Con supuesto naive Bayes= $10 * 51 - 1 = 509$

b. (*8 pts*) El clasificador de Naive Bayes asume que los atributos son independientes dado el valor de la clase objetivo. Una versión modificada y menos ingenua de este clasificador asume sólamente independencias parciales entre los atributos dado el valor de la clase. Por ejemplo, si en el problema anterior, existe una dependencias entre cada atributo a_i y a_{i-1}, tal como muestra el siguiente diagrama:



- ¿Cuál sería en este caso la ecuación para estimar $P(C|a_1, \ldots, a_n)$?

  Sol:
  $P(C|a_1, \ldots, a_n) \propto P(C)P(a_1|C) \prod_{i=2}^{n} P(a_i|C, a_{i-1})$

- ¿Cuál es el número de parámetros a estimar en este caso?

Sol:
Estimación $P(C)$ implica estimar $r_o - 1$ parámetros
Estimación $P(a_1|C)$ implica estimar r_o parámetros
Estimación $\prod_{i=2}^{n} P(a_i|C, a_{i-1})$ implica estimar $(n - 1)(2r_o)$ parámetros
TOTAL: $(r_o - 1) + r_o + (n - 1)(2r_o) = 2r_on - 1$ parámetros.

4. (*16 puntos*) **Sistemas de decisión markovianos (MDPs)**

a. (*8 pts*) Asuma un MDP con el siguiente modelo:

$P(Estado = S2|Estado = S1, A = a1) = 1{,}0$
$P(Estado = S1|Estado = S1, A = a2) = 0{,}95$
$P(Estado = S4|Estado = S1, A = a2) = 0{,}05$
$P(Estado = S1|Estado = S1, A = a3) = 1{,}0$

$P(Estado = S1|Estado = S2, A = a1) = 1{,}0$
$P(Estado = S2|Estado = S2, A = a2) = 0{,}5$
$P(Estado = S5|Estado = S2, A = a2) = 0{,}5$

$P(Estado = S3|Estado = S3, A = a1) = 1{,}0$
$P(Estado = S3|Estado = S3, A = a2) = 0{,}5$
$P(Estado = S2|Estado = S3, A = a2) = 0{,}5$
$P(Estado = S2|Estado = S3, A = a3) = 1{,}0$

$P(Estado = S4|Estado = S4, A = a1) = 1{,}0$

$P(Estado = S2|Estado = S5, A = a1) = 0{,}2$
$P(Estado = S5|Estado = S5, A = a1) = 0{,}8$
$P(Estado = S5|Estado = S5, A = a2) = 0{,}9$
$P(Estado = S6|Estado = S5, A = a2) = 0{,}1$

$P(Estado = S6|Estado = S6, A = a1) = 0{,}8$
$P(Estado = S3|Estado = S6, A = a1) = 0{,}2$
$P(Estado = S6|Estado = S6, A = a2) = 0{,}9$
$P(Estado = S5|Estado = S6, A = a2) = 0{,}1$

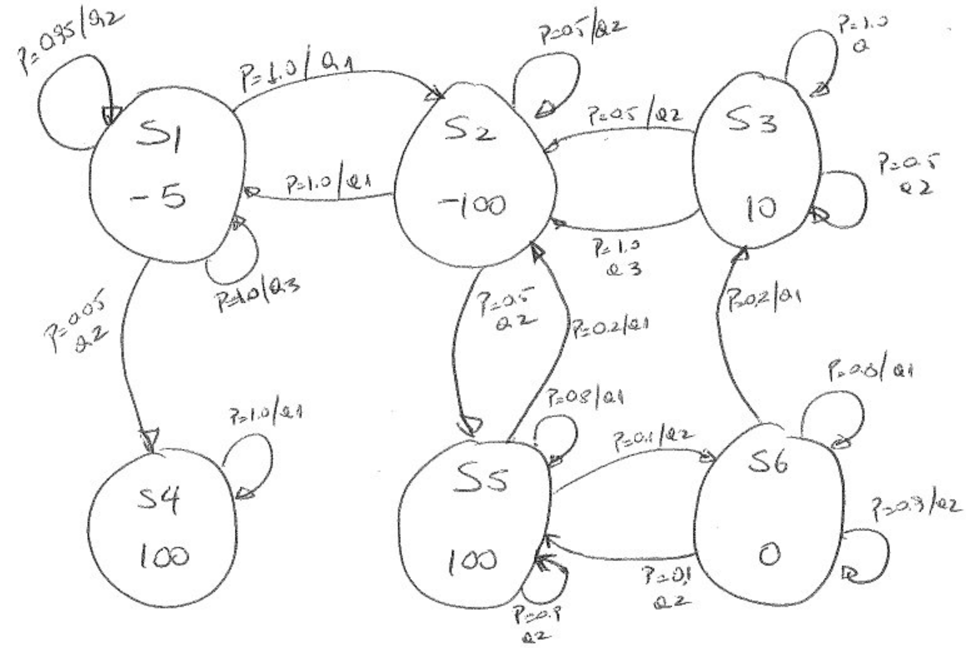
$r(S1) = -5, r(S2) = -100, r(S3) = 10, r(S4) = 100, r(S5) = 100, r(S6) = 0$

¿Cuál es la politica óptima al usar un factor de descuento $\gamma = 0{,}99$?
$\pi(S1) =?,\ \ \pi(S2) =?,\ \ \pi(S3) =?,\ \ \pi(S4) =?,\ \ \pi(S5) =?,\ \ \pi(S6) =?.$

Ayuda: note que el factor de descuento es muy cercano a 1.

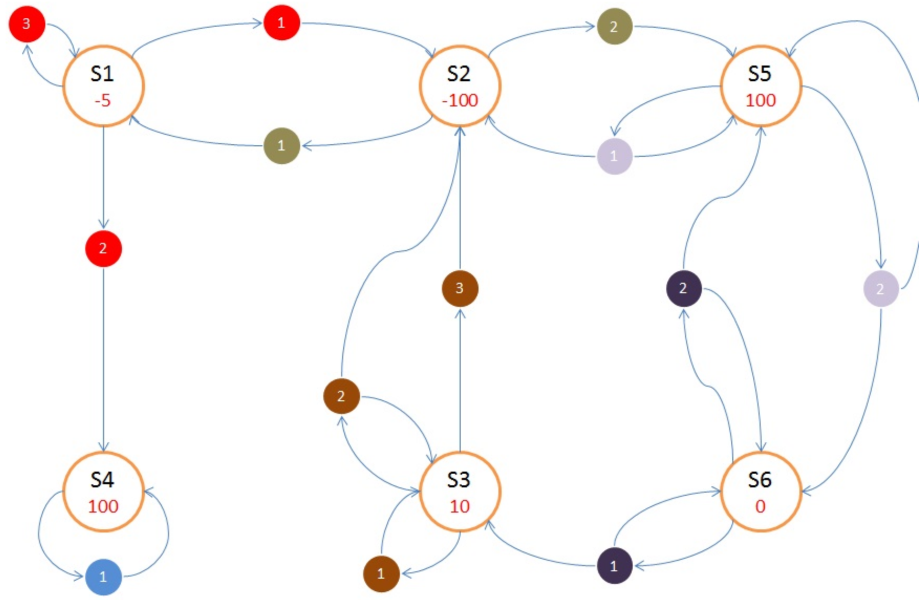
Sol:

La figura muestra gráficamente el problema:



El factor de descuento es cercano a 1, por tanto, practicamente no existe descuento por recompenzas lejanas en el tiempo. En este contexto, la recompenza en el largo plazo tiene un gran efecto en la política de acción. En el caso del problema planteado, el sistema tratará de llegar a S4 donde existe una gran recompenza y el

sistema con probabilidad 1 se mantiene en ese estado. De acuerdo a esto la política óptima prefiere acciones que facilitan la llegada más rápida al estado S4, o sea:
$\pi(S1) = a2$
$\pi(S2) = a1$
$\pi(S3) = a3$
$\pi(S4) = a1$
$\pi(S5) = a1$
$\pi(S6) = a1$



b. (*8 pts*) En cierto MDP para el estado S3 el agente puede ejecutar 3 posibles acciones. Para este caso la función de transición es la siguiente:
$P(Estado = S1 | Estado = S3, A = a1) = 1{,}0$
$P(Estado = S2 | Estado = S3, A = a2) = 1{,}0$
$P(Estado = S1 | Estado = S3, A = a3) = 0{,}4$
$P(Estado = S2 | Estado = S3, A = a3) = 0{,}6$

Si al ejecutar el algoritmo de Policy Iteration se decide ignorar en el estado S3 la posibilidad de ejecutar la acción a3, ¿Es posible que esta decisión implique que el algoritmo converja a una política de acción subóptima?. Justifique su respuesta.

Sol:

No afecta el óptimo. Esto se explica por las transiciones determinísticas desde S3 a S1 o S2 al ejecutar las acciones a1 o a2, respectivamente. Dado esto, si la solución óptima prefiriera un camino hacia S1 o S2, siempre seleccionaría las acciones a1 o a2 en lugar de a3. Por tanto, podemos estar seguros que la solución óptima no incorpora a3 para el estado S3.