

Técnicas de Inteligencia Artificial – Examen evaluación continua - 17 de enero de 2022

1.- Clasificación multiclase (4 puntos)

1.1 Dibujar la red (1 punto)

Dados items de entrada que vienen representados por 4 rasgos (features) es decir, que cada item del entrenamiento viene representado por un vector de dimensión $[1 \times 4]$ dibuja una red neuronal de 2 capas intermedias, donde la primera capa consta de 4 neuronas y la segunda de 3, sabiendo que esta red se empleará para realizar clasificación multiclase consistente en 3 clases. Decide tu la dimensión de la última capa. Inicializa los pesos con valores al azar, pero dibuja dichos valores también en el dibujo, por ejemplo $w_1=0.2$, $w_2=0.15$, etc.

1.2 Calcular el error (3 puntos)

Una vez hecho el dibujo, calcula la salida y el error empleando cross-entropy sabiendo que la clase real es la 2, es decir el vector y_{real} será 0,1,0. Presenta las operaciones (no hace falta que las resuelvas) para actualizar un peso de la segunda capa. Consulta la hoja de fórmulas.

2. Value Iteration (3 puntos)

Dada la fórmula de Bellman adaptada al Value Iteration:

$$V_{k+1} \leftarrow \max_a \sum_{s'} T(s, a, s') [R(s, a, s') + \gamma V_k(s')]$$

- Conjunto de estados: $S = \{S_0, S_1, S_2\}$
- Conjunto de acciones: $A = \{A_0, A_1\}$
- Función de transición de estados: $T : S \times A \times S \rightarrow P(S)$ donde

$T(S_0, A_0, S_0) = 0,5$
 $T(S_0, A_0, S_1) = 0$
 $T(S_0, A_0, S_2) = 0,5$
 $T(S_0, A_1, S_0) = 0$
 $T(S_0, A_1, S_1) = 0$
 $T(S_0, A_1, S_2) = 1$

$T(S_1, A_0, S_0) = 0,7$
 $T(S_1, A_0, S_1) = 0,1$
 $T(S_1, A_0, S_2) = 0,2$
 $T(S_1, A_1, S_0) = 0$
 $T(S_1, A_1, S_1) = 0,95$
 $T(S_1, A_1, S_2) = 0,05$

$T(S2, A0, S0) = 0,4$
 $T(S2, A0, S1) = 0$
 $T(S2, A0, S2) = 0,6$
 $T(S2, A1, S0) = 0,3$
 $T(S2, A1, S1) = 0,3$
 $T(S2, A1, S2) = 0,4$

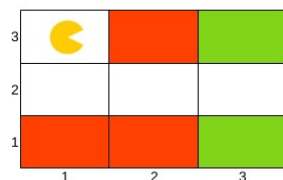
$R : S \times A \times S \rightarrow R$ donde
 $+5$ si $(s, a, s') = (S1, A0, S2)$
 -1 si $(s, a, s') = (S2, A1, S0)$
 0 en otro caso

Dibuja gráficamente el MDP correspondiente (1 punto).

Calcula el valor del estado S1 empleando el algoritmo de Value Iteration o Iteración por valor, en su versión síncrona, suponiendo que todos los estados se inicialicen a valor 0.0 (2 puntos)

3. Q-learning (3 puntos)

Dada la cuadrícula que se muestra a continuación y un agente que está tratando de aprender la política óptima, sabemos que las recompensas se consiguen solamente por realizar la acción **salir** desde uno de los estados coloreados. Al realizar esta acción, el agente pasa al estado **Final (D)** que aunque no aparece en la cuadrícula es un estado al que se puede transitar desde cualquiera de los estados coloreados aunque con distinto premio, al salir desde los estados verdes se obtiene un premio de +30 y desde los estados rojos un premio de -100. En cualquier caso, al salir el episodio terminaría. Supongamos $\gamma = 1$ y $\alpha = 0.5$ para todos los cálculos.



Sabiendo que la fórmula de Bellman adaptada para Q-learning es la siguiente:

$$\underbrace{NewQ(s, a)}_{\text{New Q value for that state and that action}} = \underbrace{Q(s, a)}_{\text{Current Q value}} + \underbrace{\alpha}_{\text{Learning Rate}} [\underbrace{R(s, a)}_{\text{Reward for taking that action at that state}} + \underbrace{\gamma}_{\text{Discount rate}} \underbrace{\max_{a'} Q'(s', a')}_{\text{Maximum expected future reward given the new s' and all possible actions at that new state}} - Q(s, a)]$$

Dados los episodios que aparecen a continuación, completa el momento en el que los siguientes valores de Q se vuelven distintos de cero por primera vez. Tu respuesta debe tener el formato (# de episodio, # de iter) donde iter # es la iteración de actualización de Q-learning en ese episodio. Si el valor Q especificado nunca pasa a ser distinto de cero, escribe nunca.

$Q((1,2), E) =$ _____
 $Q((2,2), E) =$ _____
 $Q((3,2), S) =$ _____

Episodio 1

(1,3), S, (1,2), 0
(1,2), E, (2,2), 0
(2,2), E, (3,2), 0
(3,2), N, (3,3), 0
3,3), Exit, D, +50

Episodio 2

(1,3), S, (1,2), 0
(1,2), E, (2,2), 0
(2,2), S, (2,1), 0
(2,1), Exit, D, -100

Episodio 3

(1,3), S, (1,2), 0
(1,2), E, (2,2), 0
(2,2), E, (3,2), 0
(3,2), S, (3,1), 0
(3,1), Exit, D, +30

Episodio 4

(1,3), S, (1,2), 0
(1,2), E, (2,2), 0
(2,2), E, (3,2), 0
(3,2), N, (3,3), 0
(3,3), Exit, D, +50

Episodio 5

(1,3), S, (1,2), 0
(1,2), E, (2,2), 0
(2,2), E, (3,2), 0
(3,2), S, (3,1), 0
(3,1), Exit, D, +30