Fundamentos matemáticos del aprendizaje profundo

1° cuatrimestre 2025

Práctica 5: Neuronas abstractas

Ejercicio 1. Recordemos que $\neg x$ es la negación de la variable Booleana x.

(a) Mostrar que un perceptrón puede aprender la función Booleana $y = x_1 \wedge \neg x_2$, con $x_1, x_2 \in \{0, 1\}$

(b) Misma pregunta que en (a) para la función Booleana $y = x_1 \vee \neg x_2$.

(c) Mostrar que un perceptrón con una entrada Booleana x puede aprender la negación $y = \neg x$. ¿Qué sucede si queremos usar una neurona lineal?

(d) Mostrar que un perceptrón con tres entradas Booleanas x_1, x_2, x_3 puede aprender $x_1 \wedge x_2 \wedge x_3$. ¿Qué sucede con $x_1 \vee x_2 \vee x_3$?

Ejercicio 2. Mostrar que dos conjntos finitos y linealmente separables A y B pueden ser separados por un perceptrón con pesos racionales.

Ejercicio 3. (a) Asumamos que las entradas de una neurona lineal son independientes y normalmente distribuidas, $X_i \sim N(0, \sigma_i^2)$, $i = 1, \ldots, n$. Hallas los pesos óptimos $\mathbf{w}^*$.

(b) Una variable aleatoria unidimensional con media cero, Z, es aprendida por una neurona lineal con entrada X. Asumamos que la entrada $\mathbf{X}$ y el objetivo Z son independientes. Escribir la función de costo y hallar los parámetros óptimos $\mathbf{w}^*$. Interpretar el resultado.

Ejercicio 4. Mostrar la equivalencia entre el algoritmo de regresión lineal y el de aprendizaje de una neurona lineal.

Ejercicio 5. Considerar n puntos, $P_1, \ldots, P_n$ incluidos en media circunferencia y notemos por $\mathbf{x}_1, \ldots, \mathbf{x}_n$ sus coordenadas. Un perceptrón puede aprender la mencionada media circunferencia usando el siguiente algoritmo:

1. Empezamos por una media circunferencia arbitraria determinada por su diámetro y un vector unitario $\mathbf{w}_0$. Entonces seleccionamos un punto incorrectamente clasificado P_{i_0}, i.e. un punto tal que $\mathbf{w}_0 \cdot \mathbf{x}_{i_0} < 0$. Ver Figura 1.
2. Rotemos el diámetro de manera tal que la nueva normal es $\mathbf{w}_1 = \mathbf{w}_0 + \mathbf{x}_{i_0}$. Mostrar que ahora el punto P_{i_0} es correctamente clasificado.
3. Repitiendo las dos etapas previas, construimos inductivamente una sucesión de vectores $\{\mathbf{w}_m\}_{m \geq 1}$ tal que $\mathbf{w}_{m+1} = \mathbf{w}_m + \mathbf{x}_{i_m}$, donde P_{i_m} es un punto mal clasificado en la etapa m. Mostrar que este proceso termina en un número finito de pasos.

Ejercicio 6. Modificar el algoritmo del ejercicio anterior para el caso en que los puntos $P_1, \ldots, P_n$ están incluidos en un semi espacio.

Ejercicio 7. Notemos por $\mathbf{1}_A(x)$ a la función caraterística (o indicadora) del conjunto A, es decir $\mathbf{1}_A(x) = 1$ si $x \in A$ y $\mathbf{1}_A(x) = 0$ si $x \notin A$.

(a) Mostrar que la función

$$\varphi(x_1, x_2) = \mathbf{1}_{\{x_2 > x_1 + 0,5\}}(x_1, x_2) + \mathbf{1}_{\{x_2 z x_1 + 0,5\}}(x_1, x_2)$$

implementa XOR.

(b) Mostrar que XOR puede ser implementada por una combinación lineal de dos preceptrones.

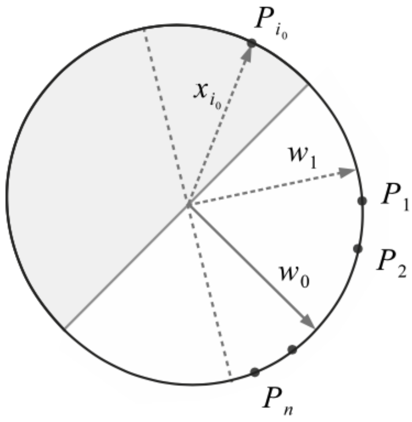

FIGURA 1. Figura para el Ejercicio 5