Anotaciones importantes sobre el XOR

*El* ***OR*** *es “linealmente Separable”, en cambio el* ***XOR*** *“no es linealmente Separable”.*

Que sea linealmente separable, significa que es posible que se resuelva el problema con 100% de efectividad solo con una recta, se pueden separar o clasificar los puntos o zonas con una sola recta.

Nos tenemos que preguntar: *¿Cuál es el menor error que podríamos esperar que tenga el problema?*

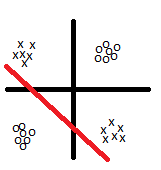
En las siguientes figuras:

-Las “x” representan 1

-Los “o” representan -1.

Para el XOR:

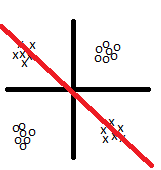
Caso 1:



Considerando que a la izquierda de la recta roja vale -1 y a la derecha de la recta roja vale 1, la recta roja representa 1/4 de exactitud respecto a “o”. En este caso, para “o” se tendría un 25% de error aprox. (podría ser también cerca de 24% o 23% de error, porque quedan esas partes blancas en los cuadrantes que harían que no sea 25% justo). En cambio para “x” se tiene un 100% de efectividad.

Por lo tanto, estaríamos teniendo en total aproximadamente un 75% de efectividad o 25% de error.

Caso 2:



Considerando que a la izquierda de la recta roja vale -1 y a la derecha de la recta roja vale 1, la recta roja representa el 50% de exactitud respecto a “o” y a “x”. Es decir, tenemos un 50% de error teniendo en cuenta ambos casos.

Por lo tanto, *el menor error que podríamos tener para el problema del XOR es 25%*, considerando el caso del *perceptrón simple*, que sería obtener algo como en el caso 1, logrando 75% de efectividad.

En caso de que se necesiten agregar más rectas, pasaríamos al caso del *perceptrón multicapa*, que es lo siguiente que veríamos (guía 2), donde la idea es ir conectando más neuronas.

Comentó el profe que, por ejemplo, con un error del 30%, se resolvería el problema en pocas épocas.

*Existen infinitas rectas que pueden ser solución al problema*.

La recta se forma con los pesos, y a su vez, los pesos son proporcionales a la taza de aprendizaje y a la diferencia entre la solución deseada y la solución obtenida.

Tenemos que tener en cuenta el valor de gamma (tasa de aprendizaje), ya que parece contradictorio, pero a veces sirve reducir el valor de gamma.

Una tasa de aprendizaje alta puede hacer que se aprenda rápido lo último que vio pero que se olvide lo anterior (esta explicado en la teoría).

**Comentarios durante la evaluación de estos dos ejercicios (2022)**

-En el archivo de test del XOR, como hay 112 (-1) y 88 (+1) en las salidas deseadas, también afecta para que nos dé cosas como 0.30 (30%) y no 0.25 (25%) como esperaríamos, porque no hay la misma cantidad de datos de cada categoría, entonces puede que la recta separe bien pero hay más datos de una de las categorías mal clasificadas o algo por el estilo.

-Para separar un archivo csv en Excel y que aparezcan los datos por columnas para contar por ejemplo la cantidad de +1 o -1, se selecciona toda la columna de datos, luego se va a “Datos” – “Texto en columna” – “Delimitados” y se elige “coma” como separador y listo.

<https://ayuda.tiendanube.com/es_ES/la-informacion-no-se-ve-separada-por-columnas-carga-masiva>

-*Más particiones significa más confianza en el resultado, no mejor resultado*.

Si uso una sola partición puede pasar que me dé por ejemplo 20% de error o que me de 60% de error, todo depende de los datos que entren en esa partición. Por eso al usar varias particiones luego sacamos una media de todas y eso nos da más confianza en el resultado.

-En el ejercicio 2, en vez de rectas tendríamos *planos* que separan las categorías (no hace falta graficarlas). El inciso “a” es no linealmente separable y el inciso “b” sí es linealmente separable, por el cambio que se hizo del -1 en el tercer renglón como indica en enunciado.