

PERCEPTRÓN Y LÓGICA DIFUSA: Computación Blanda

DANIEL GIRALDO MUÑOZ
OCTUBRE DE 2020



1 CONTENIDO

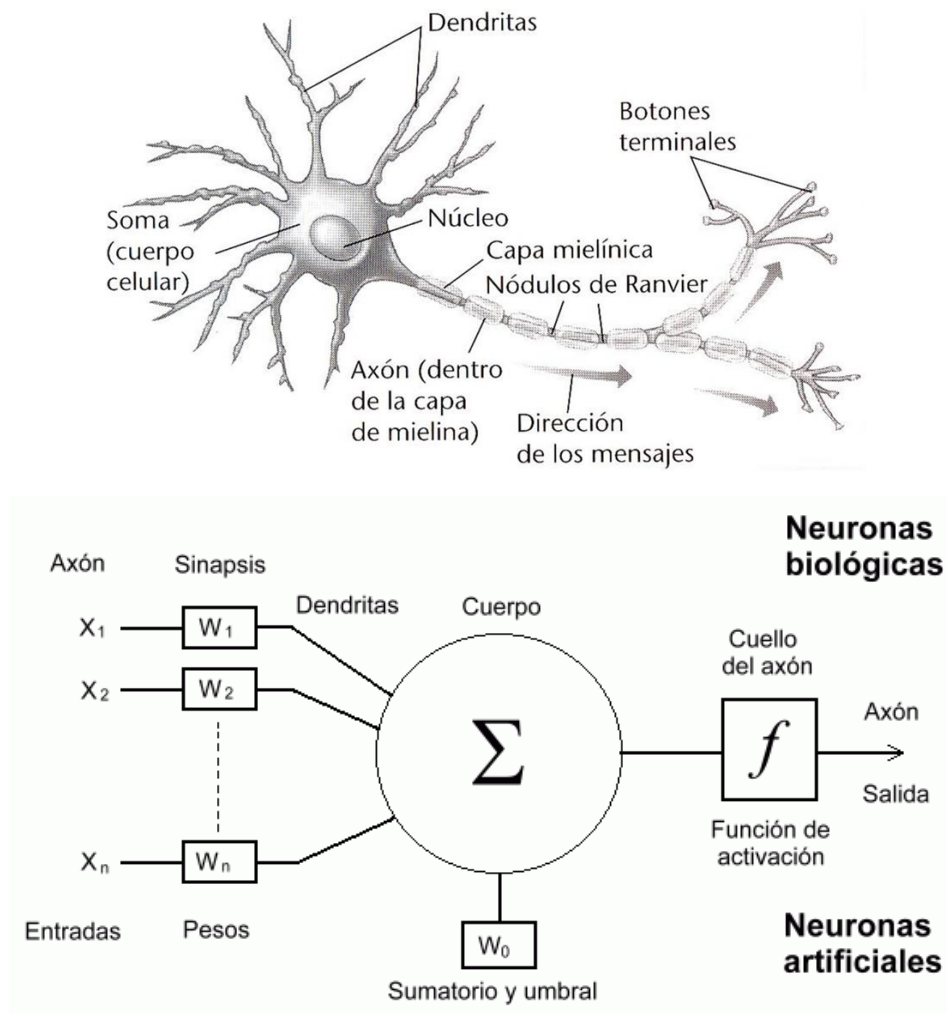
1	CONTENIDO.....	1
2	PRESENTACIÓN.....	2
3	EL PERCEPTRÓN.....	4
4	LÓGICA DIFUSA.....	12
5	CONCLUSIONES	16
6	BIBLIOGRAFÍA	17

2 PRESENTACIÓN

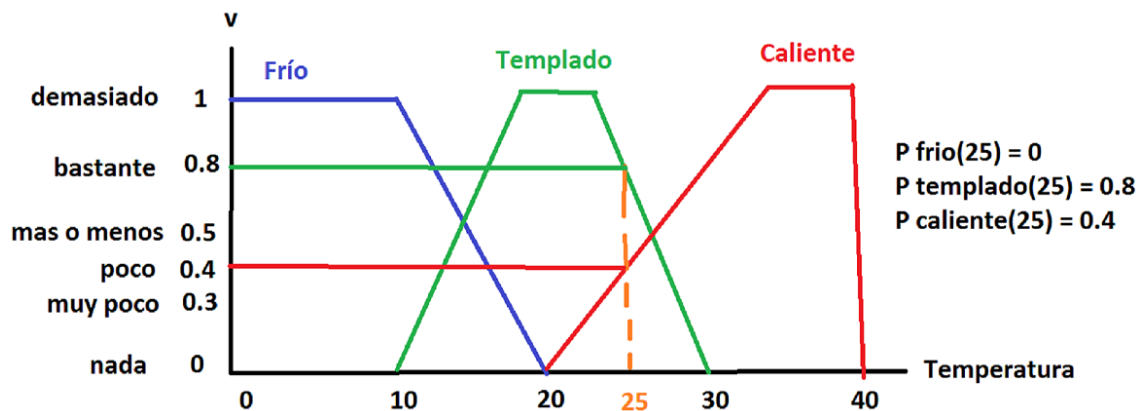
La presente monografía está orientada a la descripción de los elementos básicos de las neuronas artificiales, en particular el perceptrón, y la teoría fundamental de la lógica difusa.

En el documento se analizan los diferentes elementos que componen ambas tecnologías, mostrando las relaciones matemáticas que dan soporte a las funcionalidades tanto del perceptrón como a los factores de incertidumbre que dan sentido a la lógica difusa.

A grandes rasgos, las redes neuronales se basan en los modelos que subyacen a las redes neuronales biológicas. El siguiente diagrama adelante algunos elementos presentes en esta tecnología.



La lógica difusa se basa en la concepción de que la verdad (y la falsedad) no son absolutas. Por este motivo, todos los conceptos que concibe el ser humano tienen cierto grado de certeza, el cual se expresa fácilmente si recurrimos a un esquema como el que se ve a continuación.



En este esquema se afirma que el Frío, la sensación de Templado, y algo que es Caliente, son curvas que varían de acuerdo con la temperatura, según se ve. En el caso particular de tener una temperatura ambiente de 25 grados, dicha temperatura tendrá un valor de verdad respecto de “Caliente” de sólo 0.4. En cambio, los 25 grados representarán, en la curva de “Templado”, un valor de verdad de 0.8. Se aprecia, además, que dichos valores se relacionan, de manera bastante cercana, con frases y/o palabras que utiliza el ser humano para describir situaciones de la vida real.

En las próximas secciones se verán estas tecnologías con un mayor grado de detalle.

AUTOR: Daniel Giraldo Muñoz

1010122325

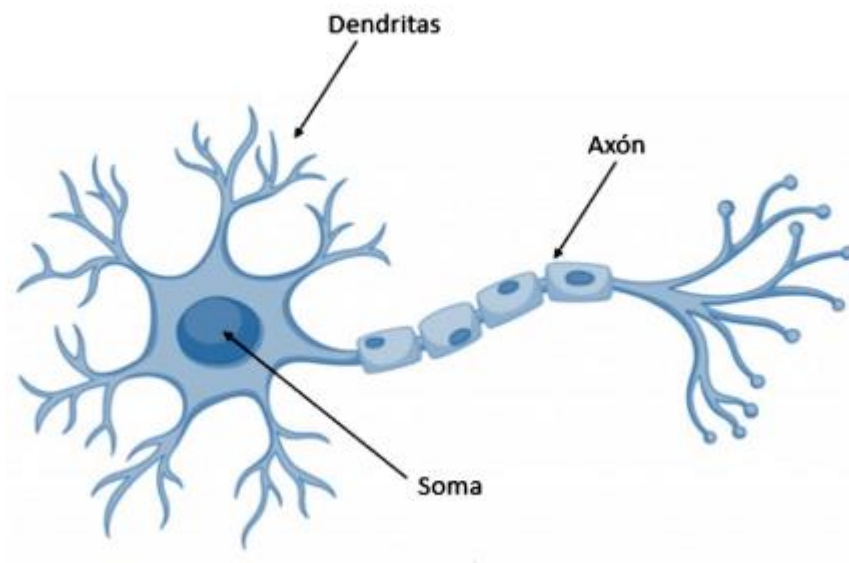
daniel.giraldo2@utp.edu.co

<https://github.com/danielgiraldo2/Computacion-Blanda.git>

3 EL PERCEPTRÓN

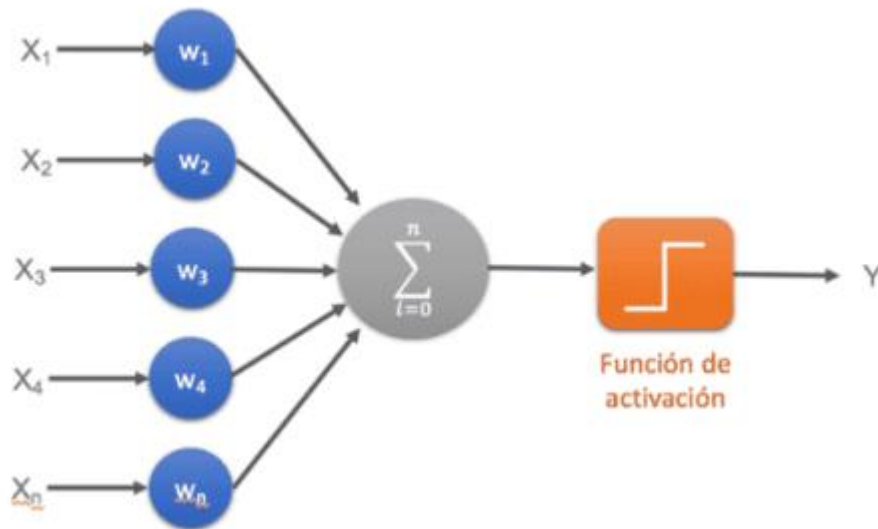
A continuación, se presenta el algoritmo básico para la conversión numérica basada en divisiones sucesivas.

Una neurona en el mundo real tiene la siguiente estructura



Siendo el Soma el núcleo, el cuerpo central que tiene unos aditamentos que permiten recoger la señales, las Dendritas son las líneas de entrada las cuales recogen las señales, el Axón es un conductor eléctrico que permite transportar esas señales eléctricas. Cuando se supera cierto nivel en el núcleo sale una señal eléctrica que se transmitirá a otras neuronas mediante los terminales del Axón.

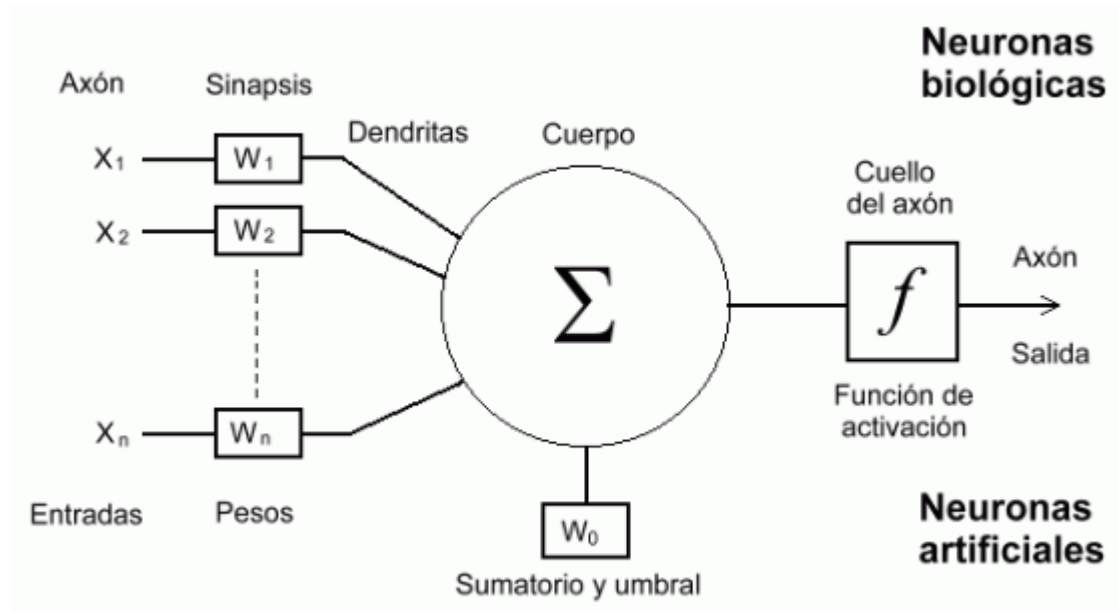
Llevando esto a un modelo artificial lo representaríamos de esta manera



Podemos observar en el lado izquierdo las respectivas entradas, las cuales serían señales eléctricas en el modelo real, pero en nuestro modelo esto serán entradas de señales binarias, en el modelo real las Dendritas tienen unos atenuadores o amplificadores de señal que hace que la neurona se proteja, si la señal aplicada es muy tensa la neurona se podría estropear, por lo tanto, lo que hacen las Dendritas es capturar una fracción de esa señal e ingresarlo a la neurona.

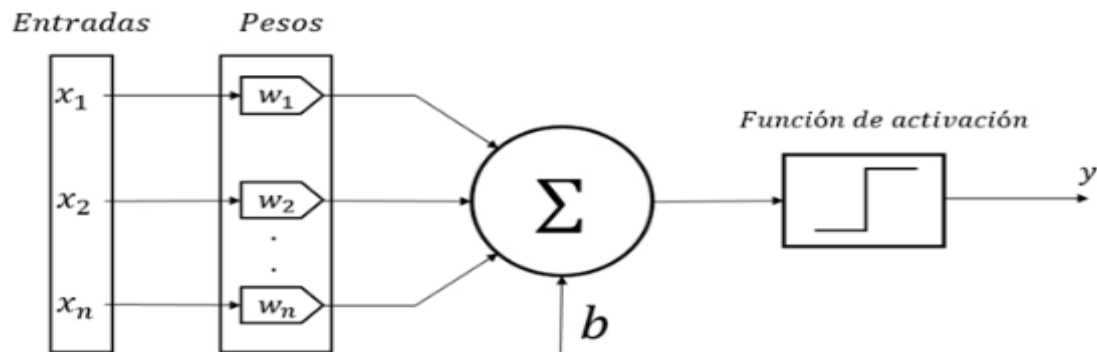
En el modelo artificial esto está representado con un concepto llamado “el peso” por lo tanto para conocer cuánta energía ingresa a la neurona, se multiplica las entradas por su peso ($X \cdot W$), realmente lo que recibe la neurona es la sumatoria de todas las señales aplicadas. Si dicha energía supera un umbral, representado en el modelo artificial como un escalón, si llega hasta el escalón y lo supera, esta neurona se activa, es decir, la neurona actúa cuando las señales de entrada están más allá del pulso que sirve como función de activación.

En el siguiente modelo podemos realizar un contraste entre el modelo real y el modelo artificial



Las entradas equivalen al Axón de otras neuronas, la ponderación de la señal, los pesos aquello que la disminuye o amplifica en la neurona real es conocida como la sinapsis, Dendritas en el mundo real pero en el modelo artificial son entradas, el cuerpo en el modelo artificial es tan solo una sumatoria de señales, pero encontramos un término llamado W_0 , es conocido como el desplazamiento, normalmente cuando la neurona se quiere activar, internamente tiene cierto repulso, este es un factor adicional aparte del umbral que se tiene que superar. Cuando la neurona supera ese umbral, la salida va a lo que en la neurona real se conoce como el cuello del Axón y en el modelo artificial se conoce como función de activación, esta función lo que hace es transferir dicha señal que ya paso el umbral y enviarlo directamente a la salida, conocida en el modelo real como un Axón de otra neurona.

A continuación, se presenta el modelo mas simple de neurona (perceptrón)



Se tienen las entradas, los pesos, la suma, la función de activación y las salidas, ilustran con esto un ejemplo real:

$$\text{UMBRAL} = U = 0.5$$

X1	X2	W1	W2	$X1*W1+X2*W2 > U$	Salida
0	0	0.3	0.3	$0*0.3 + 0*0.3 = 0.0$ NO	0

Ahora tenemos de entradas X1 y X2 con sus respectivos pesos, observando esta fila como una línea del tiempo, siendo la entrada 0 0 una posibilidad, después de tener los pesos aplicamos la ecuación anteriormente mencionada y observamos que el resultado es 0 por lo tanto estas entradas con sus señales no superan el umbral descrito en la parte superior (0.5), por lo tanto, la salida será un 0, hasta que no se supere el umbral en la salida no habrá un 1.

Ahora tenemos otras entradas X1 y X2 con sus respectivos pesos, observando esta fila como una línea del tiempo, siendo la entrada 1 1 una posibilidad diferente a la anterior

$$\text{UMBRAL} = U = 0.5$$

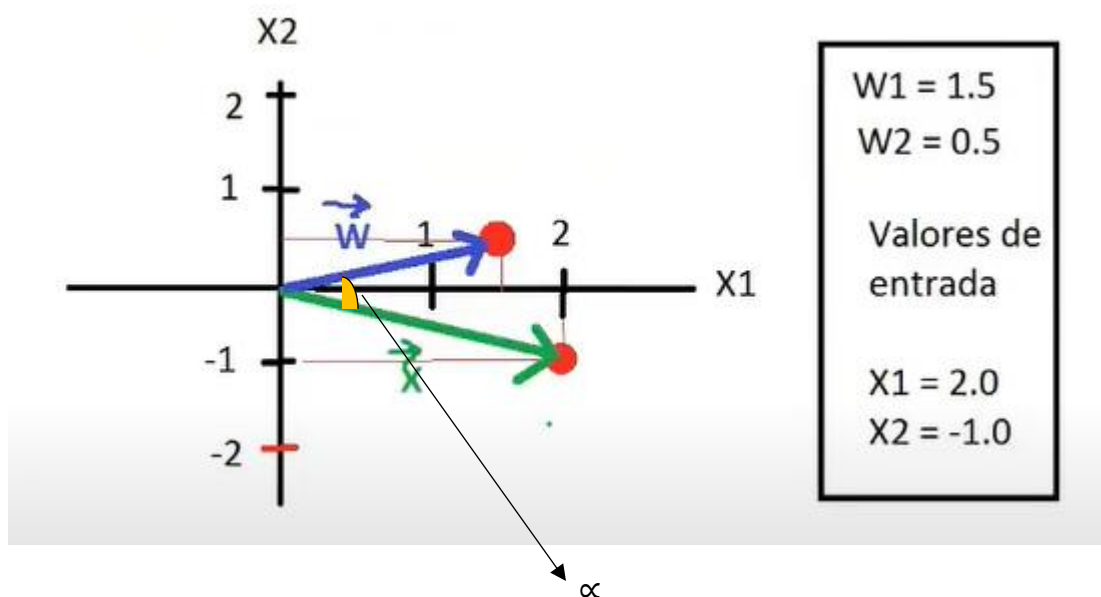
X1	X2	W1	W2	$X1*W1+X2*W2 > U$	Salida
1	1	0.3	0.3	$1*0.3 + 1*0.3 = 0.6$ SI	1

Esto nos da una sumatoria (0.6) superior a 0.5 por tanto la salida se dispara y automáticamente se convierte en 1, contrastando esto con una neurona real no sería así, ya que en una neurona real es un pulso que decae en el tiempo.

Pero ahora encontramos un dilema, que pasaría si en supuesto caso de que el valor de la ecuación de los pesos no sea ni mayor ni menor que el umbral, por lo contrario, sea igual que el umbral, entonces no se sabe cuál es la salida.

Entramos en una zona incierta (frontera), las neuronas trabajan o por encima del umbral o por debajo pero no en la frontera, pero la frontera será el secreto para entender la neurona.

Imaginemos que todo en el mundo de la neurona son espacios neuronales, por lo tanto, las entradas (X) son un vector y los pesos (W) son otro vector por lo tanto esos vectores en el espacio forman un ángulo ilustrando en el siguiente ejemplo

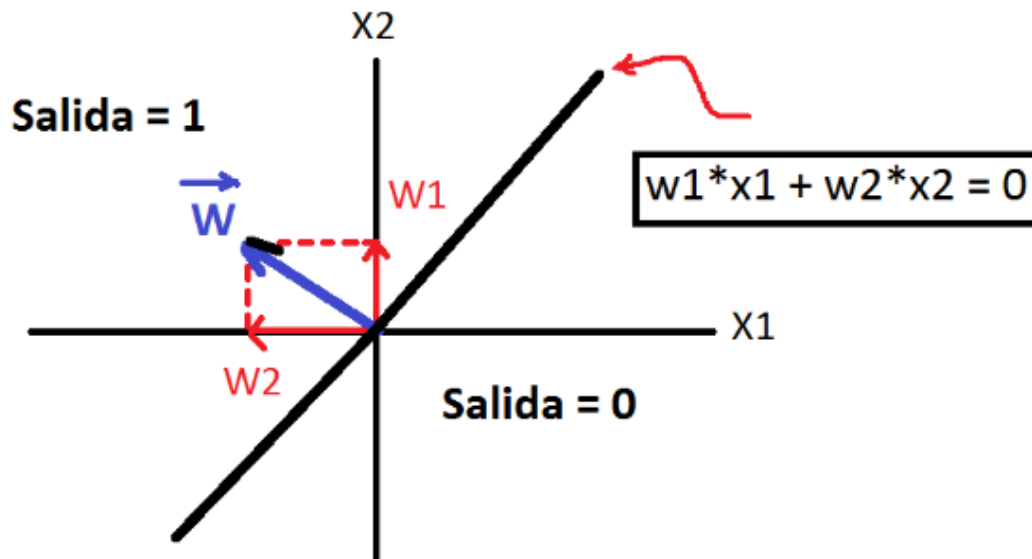


De acuerdo con esto la señal aplicada a la neurona en realidad es $\frac{\vec{w}}{|\vec{w}|} * \frac{\vec{x}}{|\vec{x}|} * \cos(\alpha)$ (la magnitud de W por la magnitud de X por el coseno del ángulo) con esto hallaremos la energía aplicada a la neurona.

Hay una condición la cual se debe cumplir para esta aplicación, la condición es que el vector W debe ser perpendicular al vector X porque el producto punto es igual a 0 significa que estamos en la frontera.

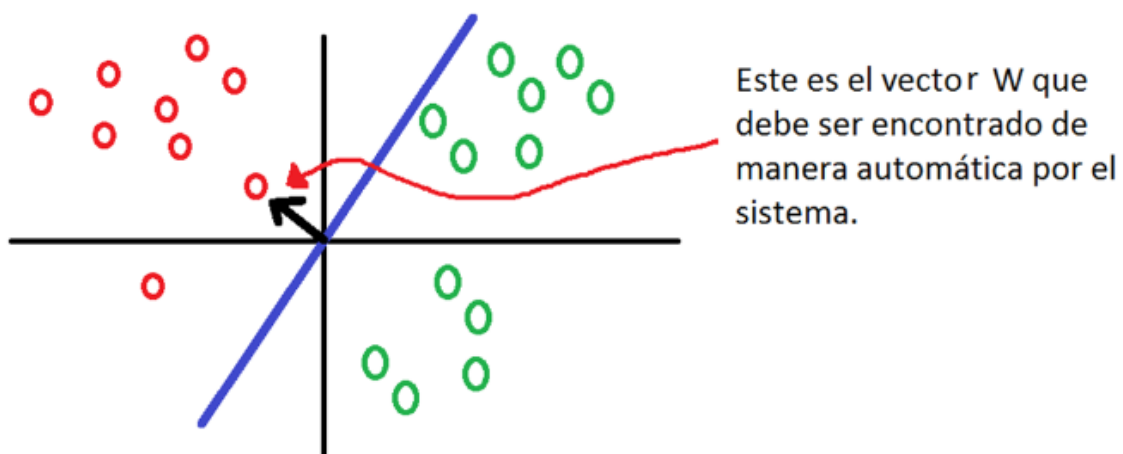
Con esta condición podremos dividir el espacio en dos zonas con una recta, la cual es la recta frontera del perceptrón, hacia donde apunte w, todos los infinitos puntos para donde

apunta W hacen que la salida valga 1, para el otro lado hace que la salida valga 0, es decir, en un lado de la recta la salida es 0 y en el otro lado la salida es 1

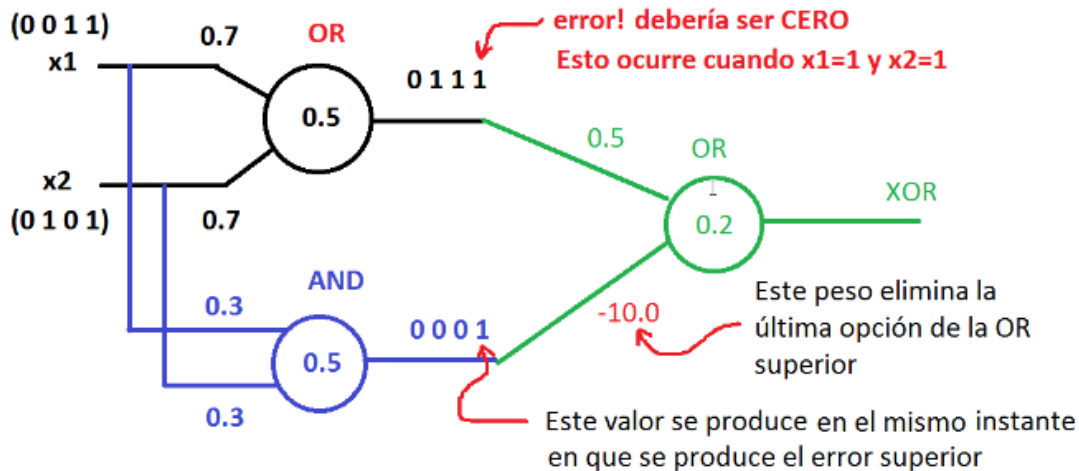


Esta recta frontera divide el espacio en dos áreas, a un lado todos los puntos que valen 1 y al otro lado todos los puntos que hacen que la salida valga 0.

Para poder calcular automáticamente la división del espacio en dos áreas de clasificación debemos modificar los pesos de forma iterativa, hasta que dichos pesos sean un vector perpendicular a la recta solución.



Esto lo hacemos mediante una compuerta XOR, pero una compuerta XOR no permite ser clasificada por una sola neurona, por lo tanto, realizamos esto mediante una compuerta OR y una compuerta AND



Tenemos la primera neurona la cual es un OR, pero tenemos el problema de que nuestro ultimo digito es un 1 y necesitamos un 0, creamos una neurona la cual es una AND para pasar ese 1 a un 0, aprovechamos que las dos neuronas coinciden en ese ultimo valor y creamos una nueva neurona la cual es una OR con un umbral de 0.2 haciendo que la neurona AND produzca un peso de -10.0 y de esta manera no pueda activar la neurona OR por lo tanto la salida se vuelve 0.

Para entrenar esta red neuronal comenzamos poniendo un punto w cualquiera que debe ser entrenado, es decir, necesitamos verificar que ese w si cumple con las expectativas de ese punto, el punto dice ser de salida 1 miramos si es menor a 90 grados, si es así, entonces cumple la condición y efectivamente es de salida 1, seguimos con otro punto el cual dice ser de salida 0 miramos si mayor a 90 grados (esto sigue siendo con el mismo w), si es así, entonces cumple la condición y efectivamente es de salida 0, de esta manera la red neuronal esta entrenada y se podrá inyectar cualquier punto y sabara ubicarlo de manera que la salida sea perfecta.

Pero qué pasa si comenzamos poniendo un punto w de tal modo que en el momento en que aparezca un punto el cual dice ser de salida 0, pero no es así, significa que el w está mal.

Debemos mover a w ya que esta mal puesto por que seguramente esta muy cerca, así que debemos alejarlo para que produzca un ángulo de 90 grados, multiplicaremos w por -1 y lo



sumamos con X de esta manera al realizar la suma vectorial, W rota pero no de una manera exagerada y justo en ese momento el punto que decía ser de salida 0 ahora si se vuelve de salida 0.

Y ahora supongamos que tenemos el mismo problema, pero en sentido contrario comenzamos poniendo un punto w de tal modo que en el momento en que aparezca un punto el cual dice ser de salida 1, pero no es así, significa que el W está mal.

Realizamos el mismo procedimiento anterior, pero con una variante Debemos mover a W ya que está mal puesto porque seguramente está muy lejos, así que debemos acercarlo para que produzca un ángulo de 90 grados, sumamos W con X de esta manera al realizar la suma vectorial, W se acerca, pero no de una manera exagerada y justo en ese momento el punto que decía ser de salida 1 ahora si se vuelve de salida 01.

Este algoritmo permite que en algún momento después de un número n de iteraciones, la recta se estabiliza y el sistema está entrenado, este es el corazón del perceptrón.

4 LÓGICA DIFUSA - INTRODUCCIÓN

La teoría base de la lógica difusa se presenta a continuación.

La lógica difusa fue investigada. Por primera vez, a mediados de los años sesenta en la universidad de Berkeley (california) por el ingeniero Lofty A.Zadeh cuando se dio cuenta de lo que el llamo principio de incompatibilidad: “conforme la complejidad de un sistema aumenta, nuestra capacidad para ser precisos y construir instrucciones sobre su comportamiento disminuye hasta el umbral más allá del cual, la preciso y el significado son características excluyentes”

Introdujo entonces el concepto de conjunto difuso bajo el que reside la idea de que los elementos sobre los que se construye el pensamiento humano no son números sino etiquetas lingüísticas.

La lógica difusa permite representar el conocimiento común, que es mayoritariamente del tipo lingüístico cualitativo y no necesariamente cuantitativo, en un lenguaje matemático a través de la teoría de conjuntos difusos y funciones características asociadas a ellos.

Permite trabajar a la vez con datos numéricos y términos lingüísticos; los términos lingüísticos son inherentemente menos precisos que los datos numéricos, pero en muchas ocasiones aportan una información más útil para el razonamiento humano.

El aspecto central de los sistemas basados en la teoría de la lógica difusa es que, a diferencia de los que se basan en la lógica clásica, tienen la capacidad de reproducir aceptablemente los modos usuales del razonamiento, considerando que la certeza de una proposición es una cuestión de grado.

Aunque la lógica difusa es conocida con este nombre desde que Zadeh la bautizó así en 1965, la idea que se esconde tras ella y sus orígenes se remontan hasta 2.500 años atrás . Los filósofos griegos, Aristóteles entre ellos, consideraban que existían ciertos grados de veracidad y falsedad y Platón ya trabajó con grados de pertenencia.

Un hito importante en el desarrollo de la lógica difusa fue establecido por Assilian y Mamdani en 1974 en el Reino Unido al desarrollar el primer controlador difuso diseñado para una máquina de vapor, pero la primera implantación real de un controlador de este tipo fue realizada en 1980 por F. En 1983 Fuji aplica la lógica difusa para el control de inyección química en plantas depuradoras de agua por primera vez en Japón y en 1987 Hitachi pone en marcha un controlador fuzzy para el control del tren-metro de Sendai, y la empresa Omron desarrolla los primeros controladores difusos comerciales. Paralelamente al desarrollo de las aplicaciones de la lógica difusa, investigadores teóricos siguen, en la

década de los ochenta, el camino iniciado por Mamdani. Así, Takagi y Sugeno desarrollan la primera aproximación para construir reglas fuzzy a partir de datos de entrenamiento, y aunque en un principio no tiene mucha repercusión, más tarde será el punto de partida para investigar la identificación de modelos fuzzy. Kosko es conocido por su contribución a los sistemas neuro-fuzzy y con sus publicaciones introdujo en la lógica difusa a muchos lectores interesados en las redes neuronales.

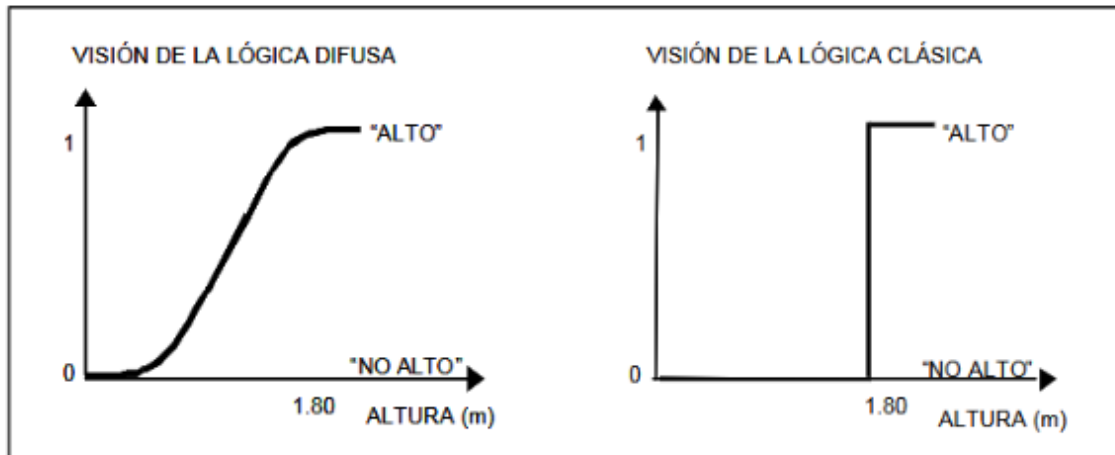
En la década de los noventa, además de las redes neuronales y los sistemas fuzzy, hacen su aparición los algoritmos genéticos. Estas tres técnicas computacionales, que pueden combinarse de múltiples maneras y se pueden considerar complementarias, son herramientas de trabajo muy potentes en el campo de los sistemas de control en la última década. En realidad, la intención original del profesor Zadeh era crear un formalismo para manipular de forma más eficiente la imprecisión y la vaguedad del razonamiento humano expresado lingüísticamente, sin embargo, causó cierta sorpresa que el éxito de la lógica borrosa llegase en el campo del control automático de procesos. Esto se debió básicamente al boom que la lógica borrosa causó en Japón, iniciado en 1987 y que alcanzó su máximo apogeo a principios de los noventa.

Desde entonces, han sido infinidad los productos lanzados al mercado que usan tecnología borrosa, muchos de ellos utilizando la etiqueta fuzzy como símbolo de calidad y prestaciones avanzadas. Estas son algunas de las muchísimas aplicaciones de la lógica difusa, que ya están funcionando en el campo de los llamados sistemas expertos. Todos estos sistemas utilizan información, esencialmente, imprecisa con el fin de lograr sus cometidos. La lógica difusa está teniendo, por lo tanto, bastante éxito en su utilización sobre los sistemas de control, aplicación que ya podría considerarse como rutinaria.

Además, según algunos de los más prestigiosos investigadores en Zadeh, para ilustrar el concepto de conjunto difuso, fue el conjunto «hombres altos». Según la teoría de la lógica clásica el conjunto “hombres altos” es un conjunto al que pertenecerían los hombres con una estatura mayor a un cierto valor, que podemos establecer en 1.80 metros, por ejemplo, y todos los hombres con una altura inferior a este valor quedarían fuera del conjunto. Así tendríamos que un hombre que mide 1.81 metros de estatura pertenecería al conjunto hombre altos, y en cambio un hombre que mida 1.79 metros de altura ya no pertenecería a ese conjunto.

El conjunto “hombres altos” es un conjunto que no tiene una frontera clara para pertenecer a él, mediante una función de transición de “alto” a “no alto” se asigna un valor de pertenencia entre 0 y 1.

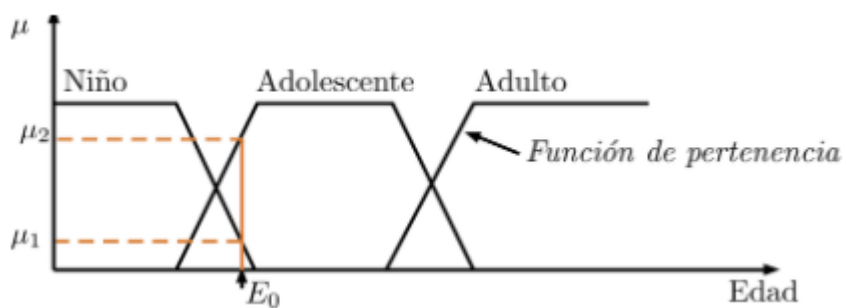
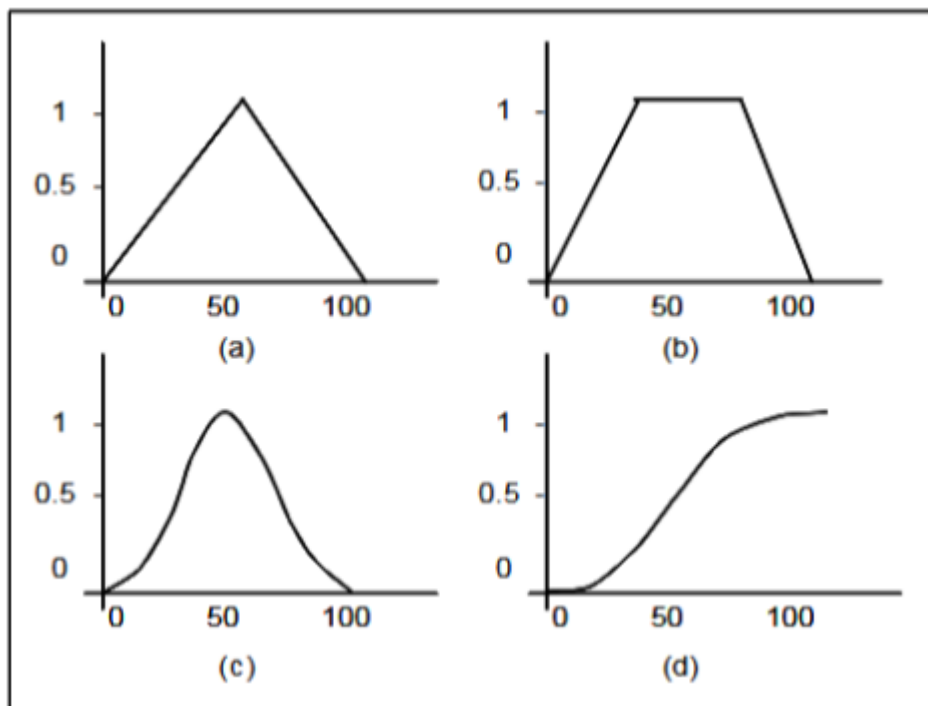
Así, por ejemplo, un hombre que mida 1.79 podría pertenecer al conjunto difuso «hombres altos» con un grado 0.8 de pertenencia, uno que mida 1 con un grado 0.85, y uno que mida 1.50 m con un grado 0.1.



La teoría clásica de conjuntos sólo contempla la pertenencia o no pertenencia de un elemento a un conjunto, sin embargo, la teoría de conjuntos difusos contempla la pertenencia parcial de un elemento a un conjunto, es decir, cada elemento presenta un grado de pertenencia a un conjunto difuso que puede tomar cualquier valor entre 0 y 1.

La función característica proporciona una medida del grado de similitud de un elemento de U con el conjunto difuso. La forma de la función característica utilizada, depende del criterio aplicado en la resolución de cada problema y variará en función de la cultura, geografía, época o punto de vista del usuario, la única condición que debe cumplir una función característica es que tome valores entre 0 y 1, con continuidad.

Conceptualmente existen dos aproximaciones para determinar la función característica asociada a un conjunto: la primera aproximación está basada en el conocimiento humano de los expertos, y la segunda aproximación es utilizar una colección de datos para diseñar la función.



5 CONCLUSIONES

El desarrollo de las temáticas elaboradas en clase utilizando las redes neuronales y la lógica difusa resultan ser unas herramientas potentes y eficaces a la hora de tener un aprendizaje supervisado, por una parte el perceptrón nos permite y nos otorga una amplia capacidad para de aprender a realizar tareas basadas en un entrenamiento o en una experiencia inicial, puede crear su propia organización o representación de la información que recibe y no solo esto , además puede ser tolerante a fallos ya que gracias a su extendida red una destrucción parcial de esta conduce a una degradación parcial de su estructura mas no a su pérdida total.

Es una herramienta potente ya que estos cálculos neuronales pueden ser realizados en paralelo, presenta grandes ventajas con respecto a otros modelos típicos de solución de problemas de Ingeniería, una de ellas es su inspiración en modelos biológicos del funcionamiento del cerebro, lo que facilita su estudio debido a las analogías que pueden introducirse para su análisis.

son una teoría relativamente nueva que junto a otras técnicas de inteligencia artificial ha generado soluciones muy confiables a problemas de Ingeniería, los cuales a pesar de poder ser solucionados por métodos tradicionales, encuentran en las redes neuronales una alternativa fácil de implementar y altamente segura, el perceptrón es una red que puede implementarse exitosamente para resolver problemas de clasificación de patrones que sean linealmente separables, la red responderá mejor entre más sencillos sean los patrones que debe clasificar. A pesar de que cuenta con serias limitaciones, esta red conserva su importancia ya que sirvió como inspiración para otros tipos de redes, como por ejemplo las redes multicapa.

Por otro lado, utilizando la lógica difusa encontramos su facilidad de implementación, permiten muy buenos resultados en procesos no lineales y de difícil modelación ya que su funcionamiento es similar al comportamiento del ser humano.

Es una forma rápida y económica de resolver un problema ya que no se necesita conocer el modelo matemático que rige su funcionamiento.

La lógica difusa admite verdades y falsedades parciales, aceptar ese tipo de premisas exige una nueva forma de ver los problemas, esta junto con otras herramientas abren las puertas al tratamiento de fenómenos y a sistemas que resultaban difíciles de definir y en ciertos casos estas tareas eran casi imposibles mediante los modelos matemáticos convencionales.



6 BIBLIOGRAFÍA

<https://repl.it>

<https://es.slideshare.net/Daniela2631/redes-neuronales-1508349>

<https://docplayer.es/39018249-la-teoria-de-redes-neuronales-artificiales-presenta-grandes-ventajas-con.html>

<https://docplayer.es/39018249-la-teoria-de-redes-neuronales-artificiales-presenta-grandes-ventajas-con.html>

<https://es.slideshare.net/MichelleDiaz6/logica-difusa>

<https://sites.google.com/site/logicadifusaingindustrialpaita/logica-difusa>

[Diapositivas vistas en clase]