

PERCEPTRÓN Y LÓGICA DIFUSA: Computación Blanda

STEVEN MEDINA GONZALEZ – ALDAHIR ROJAS LANCHEROS
OCTUBRE DE 2020



1 CONTENIDO

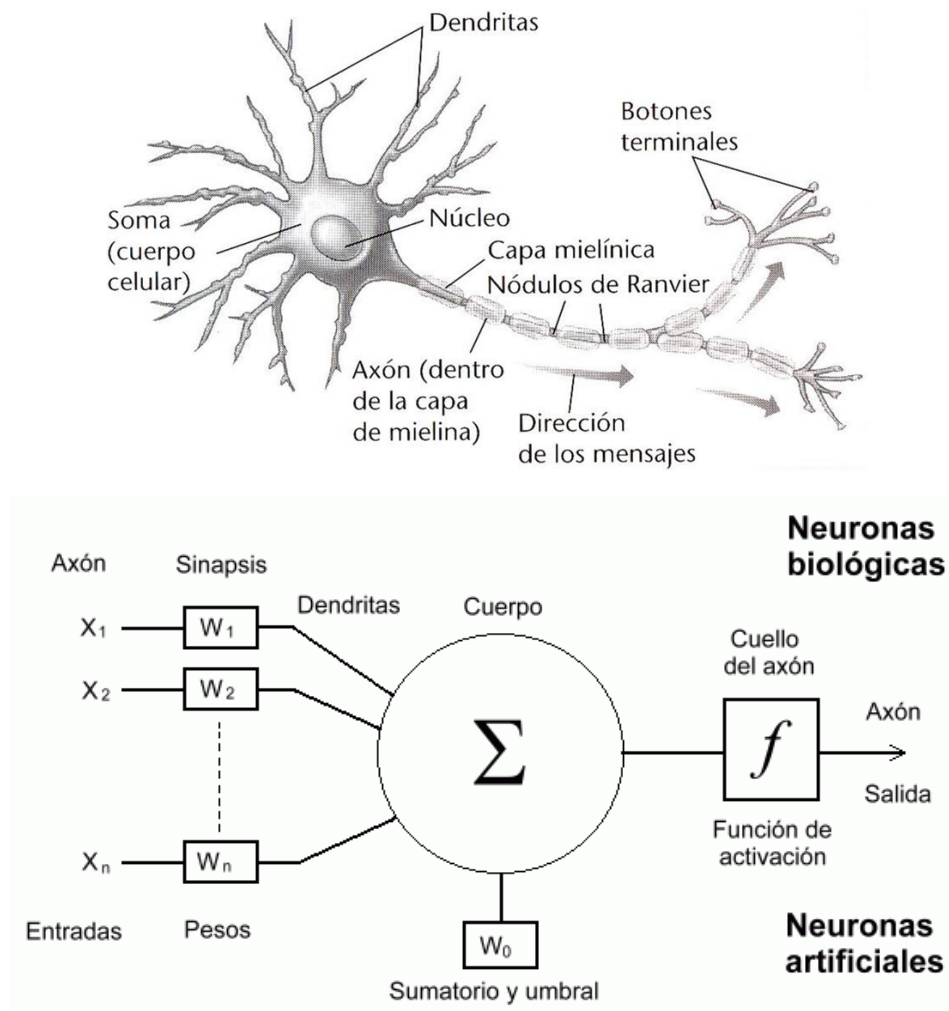
1	CONTENIDO.....	1
2	PRESENTACIÓN.....	2
3	EL PERCEPTRÓN.....	4
3.1	Arquitectura de un perceptrón	4
3.2	Funcionamiento de un perceptrón	5
4	LÓGICA DIFUSA - INTRODUCCIÓN.....	8
5	CONCLUSIONES	12
6	BIBLIOGRAFÍA.....	13

2 PRESENTACIÓN

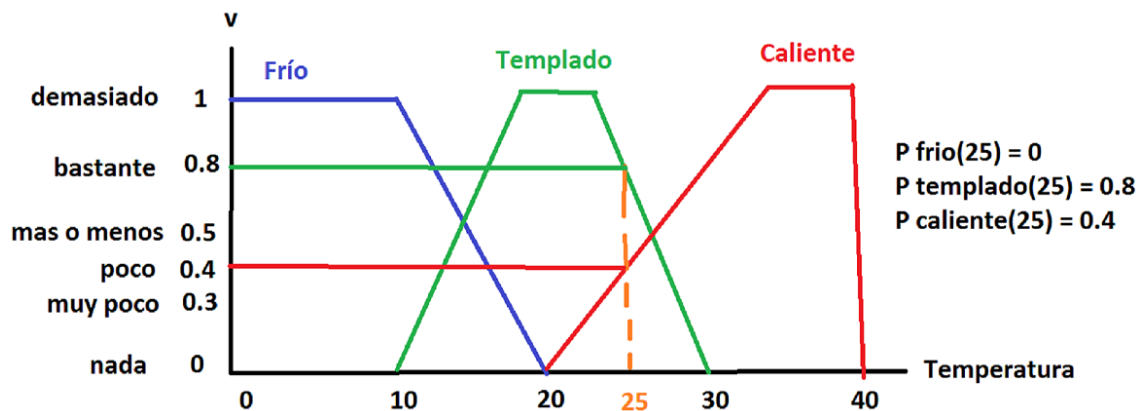
La presente monografía está orientada a la descripción de los elementos básicos de las neuronas artificiales, en particular el perceptrón, y la teoría fundamental de la lógica difusa.

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.

3 EL PERCEPTRÓN

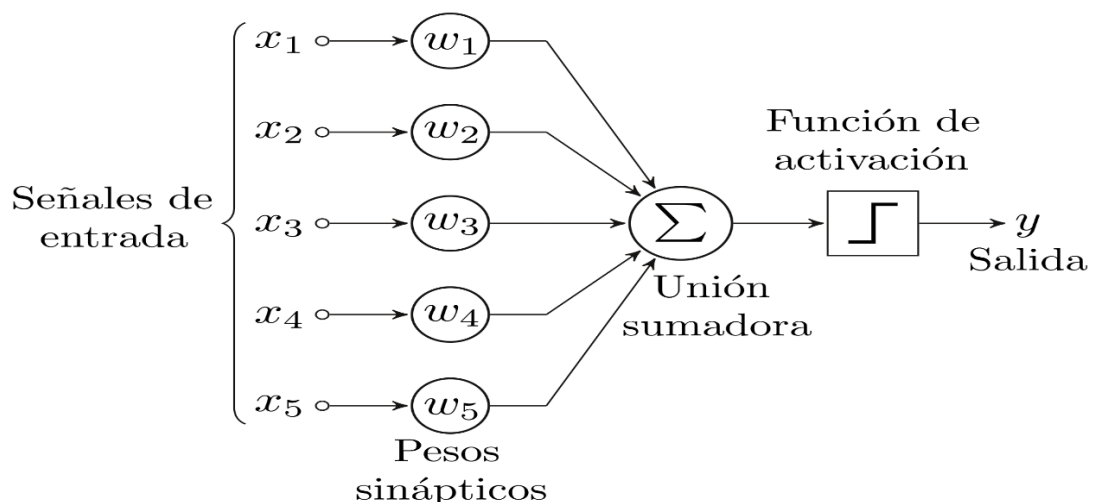
Entre las décadas de 1950 y 1960 el científico Frank Rosenblatt, inspirado en el trabajo de Warren McCulloch y Walter Pitts creó el Perceptrón, la unidad desde donde nacería y se potenciarían las redes neuronales artificiales.

El perceptrón es la red neuronal más básica que existe de aprendizaje supervisado sus principales usos son decisiones binarias sencillas, o para crear funciones lógicas como OR, AND.

El modelo biológico más simple de un perceptrón es una neurona y viceversa, como ya lo vimos en la presentación. Es decir, el modelo matemático más simple de una neurona es un perceptrón. La neurona es una célula especializada y caracterizada por poseer una cantidad indefinida de canales de entrada llamados dendritas y un canal de salida llamado axón. Las dendritas operan como sensores que recogen información de la región donde se hallan y la derivan hacia el cuerpo de la neurona que reacciona mediante una sinapsis que envía una respuesta hacia el cerebro, esto en el caso de los seres vivos.

Una neurona sola y aislada carece de razón de ser. Su labor especializada se torna valiosa en la medida en que se asocia a otras neuronas, formando una red. Normalmente, el axón de una neurona entrega su información como "señal de entrada" a una dendrita de otra neurona y así sucesivamente. El perceptrón que capta la señal en adelante se extiende formando una red de neuronas, sean éstas biológicas o de sustrato semiconductor (compuertas lógicas).

3.1 ARQUITECTURA DE UN PERCEPTRÓN



Señales de entrada x_1, x_2, \dots, x_n

Representan las entradas de la red neuronal.

Pesos sinápticos w_1, w_2, \dots, w_n

Cada entrada tiene un peso que se va ajustando de forma automática a medida que la red neuronal va aprendiendo.

Unión sumadora Σ

Realiza el sumatorio de todas las entradas ponderadas por sus pesos.

Función de activación F

Se encarga de mantener el conjunto de valores de salida en un rango determinado, normalmente $(0,1)$ o $(-1,1)$

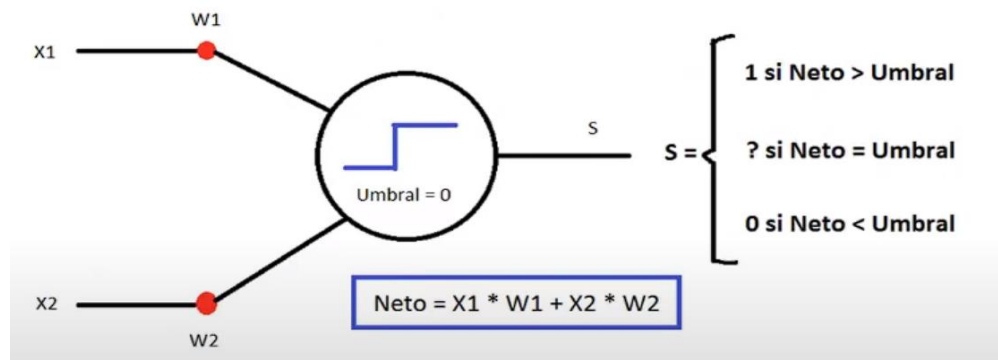
Existen diferentes funciones de activación que cumplen este objetivo, la más habitual es la función sigmoide.

Salida Y

Representa el valor resultante tras pasar por la red neuronal.

3.2 FUNCIONAMIENTO DE UN PERCEPTRÓN

El funcionamiento del perceptrón es muy sencillo, simplemente lee los valores de entrada justo después de que se hace su operación con los pesos, los cuales sirven para amplificar o reducir dependiendo de cómo se construya la neurona, suma todas y cada una de las entradas donde también se encuentra el umbral, el cual representa el grado de inhibición de la neurona, es un término constante que no depende del valor que tome la entrada.



Lo anterior es un pequeño ejemplo de como dependiendo del valor neto que se genere al sumar las entradas, el umbral permite o no una salida, dependiendo de que valor tenga este a comparación del neto, finalmente, el resultado lo introduce en una función de activación que genera el resultado final.

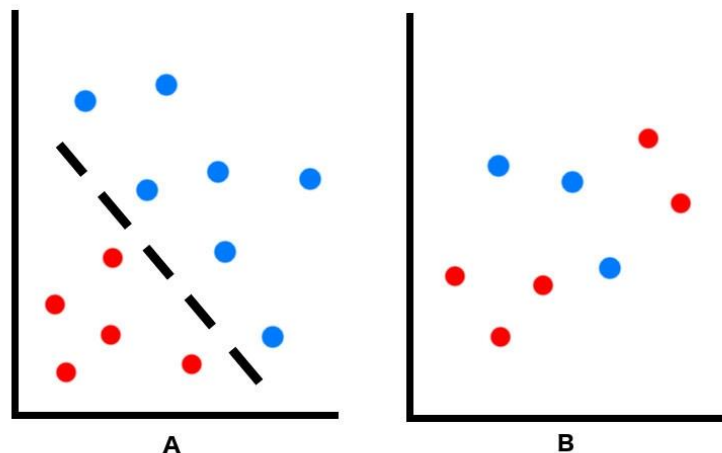
El entrenamiento del perceptrón no es más que determinar los pesos sinápticos y el umbral que mejor hagan que la entrada se ajuste a la salida. Para la determinación de estas variables, se sigue un proceso adaptativo. El proceso comienza con valores aleatorios y se van modificando estos valores según la diferencia entre los valores deseados y los calculados por la red.

En resumen, el perceptrón aprende de manera iterativa siguiendo estos pasos:

1. Inicializar pesos y umbrales
2. Bucle: hasta resultado de pesos sea aceptable
3. Bucle: para todos los ejemplos
 - Leer valores de entrada
 - Calcular error
 - Actualizar pesos según el error
 - Actualizar pesos de entradas
 - Actualizar el umbral

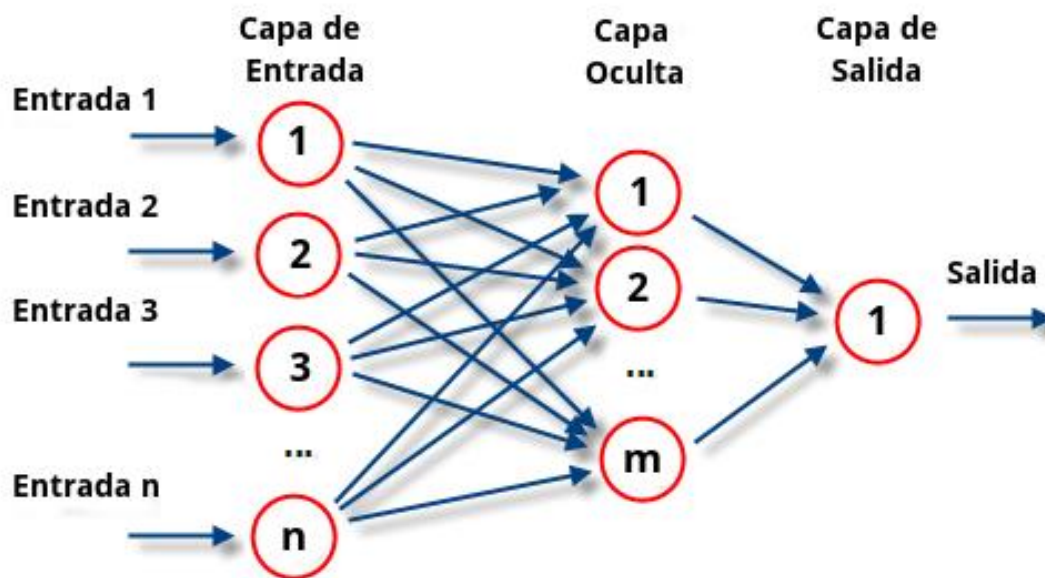
Nota: Solo es capaz de representar funciones lineales debido a que no dispone de capas ocultas como por ejemplo el perceptrón multicapa.

El perceptrón simple solo es eficaz con datos que pueden ser separables linealmente como puedes ver en la siguiente imagen:



- A: los resultados pueden ser separados linealmente.
- B: los resultados no pueden ser separados linealmente y requiere más de una línea para su separación. Para este tipo de salidas no puede ser aprendida por un perceptrón simple.

Para este tipo de problemas se usa el perceptrón multicapa, el cual es una red neuronal artificial formada por múltiples capas, esto le permite resolver problemas que no son linealmente separables, la salida de una neurona está determinada por la suma ponderada con las salidas de las neuronas de la capa anterior y una función de activación.



Las capas pueden clasificarse en tres tipos:

- Capa de entrada: Constituida por aquellas neuronas que introducen los patrones de entrada en la red. En estas neuronas no se produce procesamiento.
- Capas ocultas: Formada por aquellas neuronas cuyas entradas provienen de capas anteriores y cuyas salidas pasan a neuronas de capas posteriores.
- Capa de salida: Neuronas cuyos valores de salida se corresponden con las salidas de toda la red.

Pero esto es algo aparte de lo básico que queremos enseñar, por ello proseguiremos con el siguiente tema.

4 LÓGICA DIFUSA - INTRODUCCIÓN

La lógica difusa fue investigada, por primera vez, a mediados de los años sesenta en la Universidad de Berkeley (California) por el ingeniero Lotfy A. Zadeh cuando se dio cuenta de lo que él llamó 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 precisión y el significado son características excluyentes”.

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.

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 [46]. Los filósofos griegos, Aristóteles entre ellos, consideraban que existían ciertos grados de veracidad y falsedad y Platón ya trabajó con grados de pertenencia. El término borroso aplicado a la lógica y a la teoría de conjuntos y sistemas procede de la expresión fuzzy sets (conjuntos borrosos) acuñada por Lofti A. Zadeh, brillante ingeniero eléctrico iraní nacionalizado en Estados Unidos, profesor en las más prestigiosas universidades norteamericanas y doctor honoris causa de varias instituciones académicas.

Aunque en un principio la lógica difusa encontró una fuerte resistencia entre la comunidad científica, algunos investigadores se convirtieron en seguidores de las teorías de Zadeh y mientras él siguió ampliando y asentando los fundamentos de la teoría de conjuntos difusos estos investigadores exploraron estas nuevas teorías durante la década posterior a su nacimiento.

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.L. Smidth & Co. en una planta cementera en Dinamarca. 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.

Otro de los factores que contribuye a seguir con la investigación en este campo es el creciente interés en las redes neuronales y su similitud con los sistemas fuzzy; la tendencia es buscar vías de relación entre las dos técnicas y los resultados son los llamados neuro-fuzzy systems, sistemas fuzzy que usan métodos de aprendizaje basados en redes neuronales para identificar y optimizar

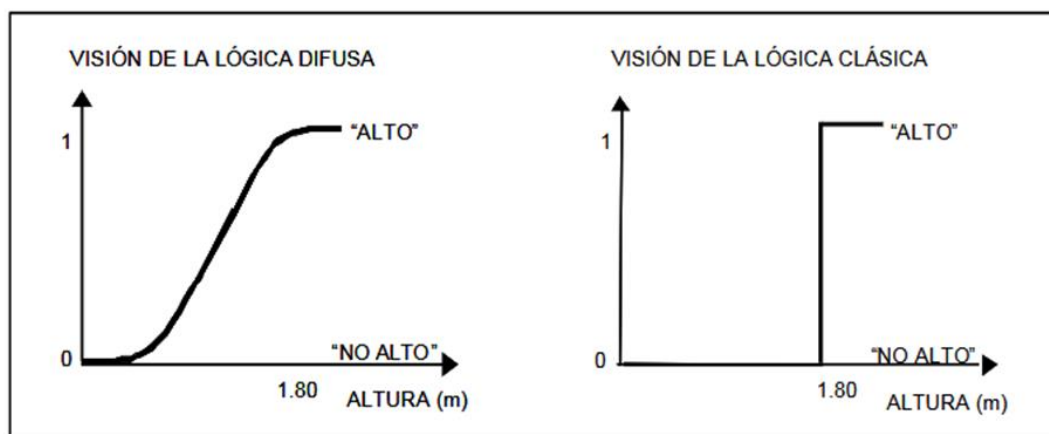
sus parámetros. B. Kosko es conocido por su contribución a los sistemas neuro-fuzzy y con sus publicaciones introdujo en la lógica difusa a muchos lectores interesados en las redes neuronales.

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 múltiple maneras y se pueden considerar complementarias, son herramientas de trabajo muy potentes en el campo de los sistemas de control en la última década.

Se investiga en áreas como el reconocimiento de patrones visuales o la identificación de segmentos de ADN, por mencionar dos ejemplos. Además, según algunos de los más prestigiosos investigadores en Internet, parece que el futuro para abordar la ingente cantidad de datos, recuperar la información, controlar y gestionar la red, pasa por el uso de las tecnologías borrosas. Esta intuición parece ser que coincide con la nueva orientación que, según el profesor Zadeh, debe seguir la lógica borrosa. Prueba de ello fue la celebración del primer encuentro sobre lógica borrosa e internet en el año 2001 (FLINT 2001) en la universidad de Berkeley organizado por el propio Zadeh.

El primer ejemplo utilizado por Lofti A. 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. Sin embargo, no parece muy lógico decir que un hombre es alto y otro no lo es cuando su altura difiere en dos centímetros.

El enfoque de la lógica difusa considera que el conjunto “hombres altos” es un conjunto que no tiene una frontera clara para pertenecer o no pertenecer a él: mediante una función que define la transición de “alto” a “no alto” se asigna a cada valor de altura un grado de pertenencia al conjunto, 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 1.81 con un grado 0.85, y uno que mida 1.50 m con un grado 0.1.



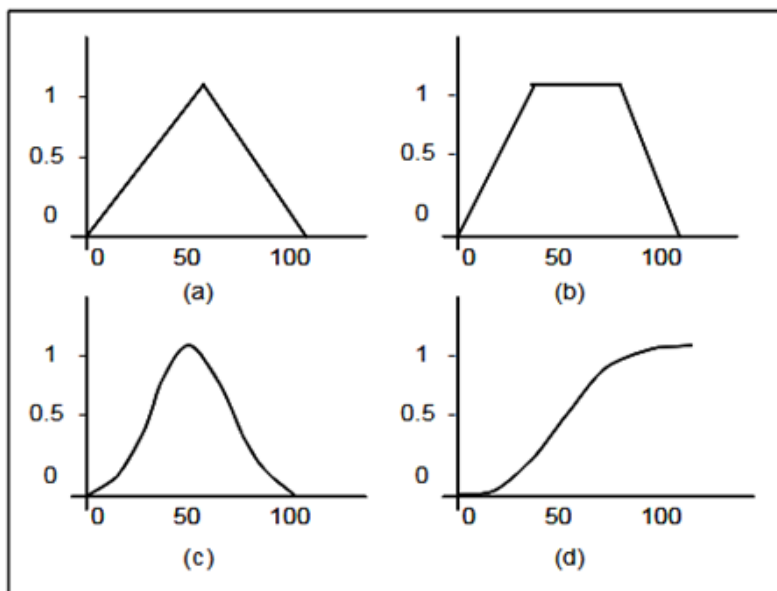
Un conjunto difuso en el universo de discurso U se caracteriza por una función de pertenencia $\mu_A(x)$ que toma valores en el intervalo $[0,1]$, y puede representarse como un conjunto de pares ordenados de un elemento x , y su valor de pertenencia al conjunto

$$A = \{(x, \mu_A(x)) | x \in U\}$$

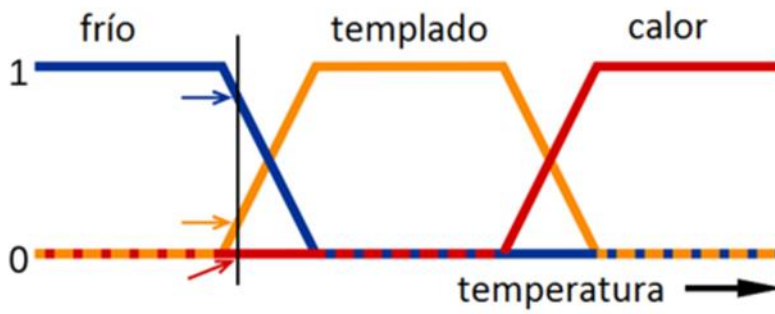
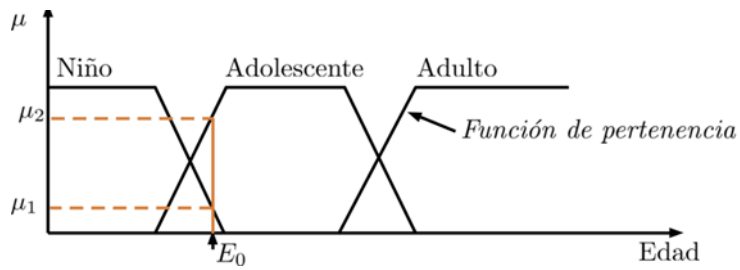
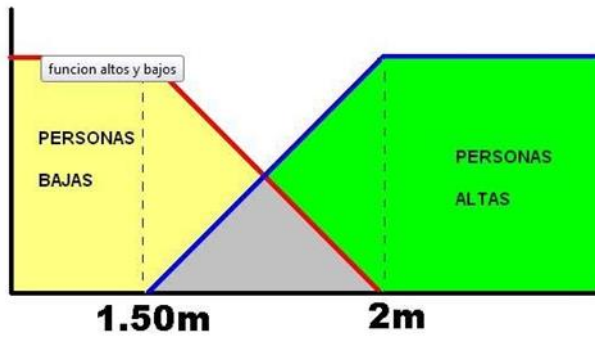
La función característica proporciona una medida del grado de similitud de un elemento de U con el conjunto difuso. La forma de la función característica utilizada, depende del criterio aplicado en la resolución de cada problema y variará en función de la cultura, geografía, época o punto de vista del usuario.

La única condición que debe cumplir una función característica es que tome valores entre 0 y 1, con continuidad. Las funciones características más comúnmente utilizadas por su simplicidad matemática y sus manejabilidades son: triangular, trapezoidal, gaussiana, sigmoideal, gamma, pi, campana etc...

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.



Otros ejemplos pueden ser:



5 CONCLUSIONES

El desarrollo de las temáticas elaboradas en clase utilizando el lenguaje JavaScript prueba ser un mecanismo de gran valor para el aprendizaje de los conceptos básicos de la materia.

Después de analizar más detenidamente el perceptrón pudimos entender la importancia que tuvo en los inicios de las redes neuronales artificiales, siendo el primer y más básico modelo en este campo, dando así un gran salto y dando pie en lo que conocemos hoy en día sobre inteligencia artificial, además de también ser simple, intuitiva pero poderosa, haciendo posible algunas implementaciones interesantes para lo que podemos poner en práctica dentro de la carrera y el campo laboral.

La lógica difusa se adapta mejor al mundo real en el que vivimos, e incluso puede comprender y funcionar con nuestras expresiones, del tipo “hace mucho calor”, “no es muy alto”, “el ritmo del corazón está un poco acelerado”, etc. La clave de esta adaptación al lenguaje se basa en comprender los cuantificadores de cualidad para nuestras inferencias (en los ejemplos anteriores, “mucho”, “muy”, y “un poco”).

Los sistemas basados en lógica difusa imitan la forma en que toman decisiones los humanos, con la ventaja de ser mucho más rápidos. En Inteligencia artificial, la lógica difusa, o lógica borrosa se utiliza para la resolución de una variedad de problemas, principalmente los relacionados con control de procesos industriales complejos y sistemas de decisión en general, la resolución y la comprensión de datos. Los sistemas de lógica difusa están también muy extendidos en la tecnología cotidiana, por ejemplo en cámaras digitales, sistemas de aire acondicionado, lavadoras, etc.



6 BIBLIOGRAFÍA

- [1] <http://mercurio.ugr.es/pedro/tutoriales/gaicm1/presentacion.pdf>
- [2] <https://www.diegocalvo.es/perceptron/>
- [3] <https://www.aprendemachinelearning.com/breve-historia-de-las-redes-neuronales-artificiales/#:~:text=Entre%20las%20d%C3%A9cadas%20de%201950,produce%20una%20s%C3%B3la%20salida%20binaria.>
- [4] <https://es.wikipedia.org/wiki/Perceptr%C3%B3n>
- [5] https://es.wikipedia.org/wiki/Perceptr%C3%B3n_multicapa
- [6] https://es.wikipedia.org/wiki/L%C3%B3gica_difusa