2020

PERCEPTRÓN Y LÓGICA DIFUSA: Computación Blanda

KAREN POSADA MUÑOZ Y JUAN CAMILO LOPERA MARTÍNEZ
OCTUBRE DE 2020

UTP | Pereira

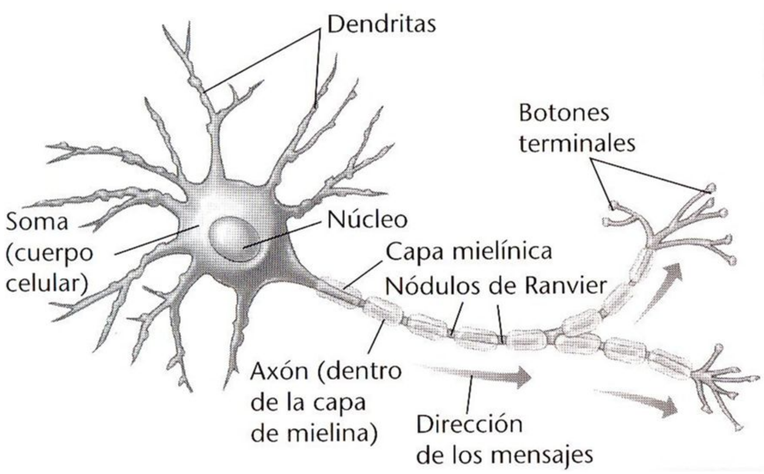
1 CONTENIDO

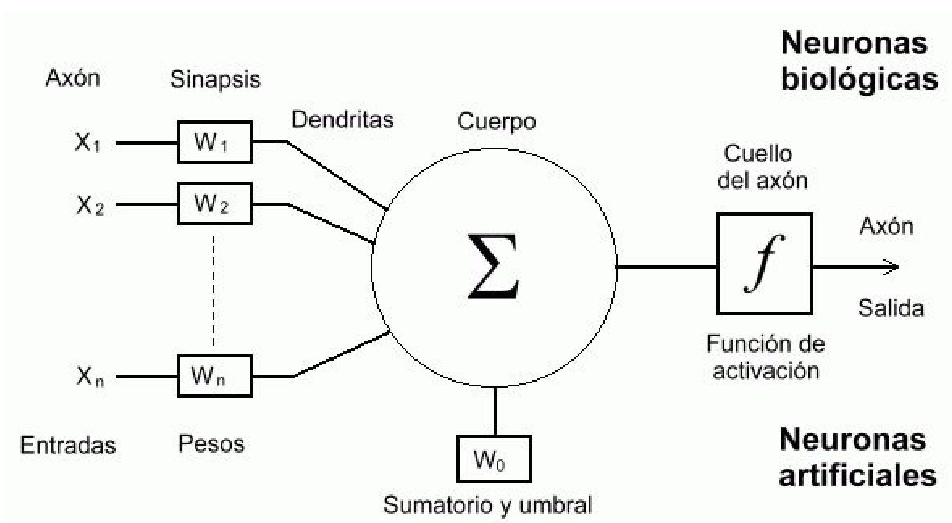
2 PRESENTACIÓN

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.

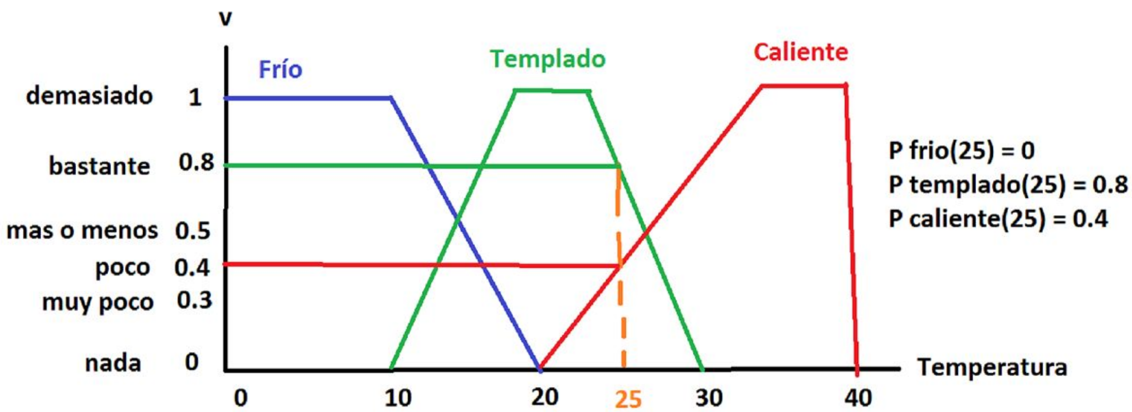
En el documento se analizan los diferentes elementos que componen ambas tecnologías, mostrando las relaciones matemáticas que dan soporte a las funcionalidades tanto del perceptrón como a los factores de incertidumbre que dan sentido a la lógica difusa.

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.

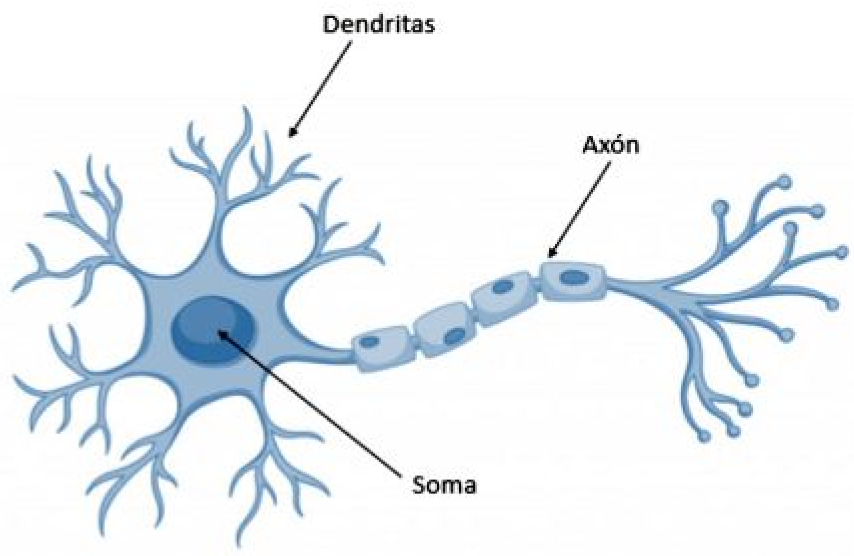
AUTOR: Karen Posada Muñoz - 1088039527 - karen.posada@utp.edu.co - https://github.com/PosadaKaren/Computacion-Blanda

Juan Camilo Lopera Martínez - 1010130512 - juancamilo.lopera@utp.edu.co - https://github.com/loperamartinez/ComputacionBlandaGR1

3 EL PERCEPTRÓN

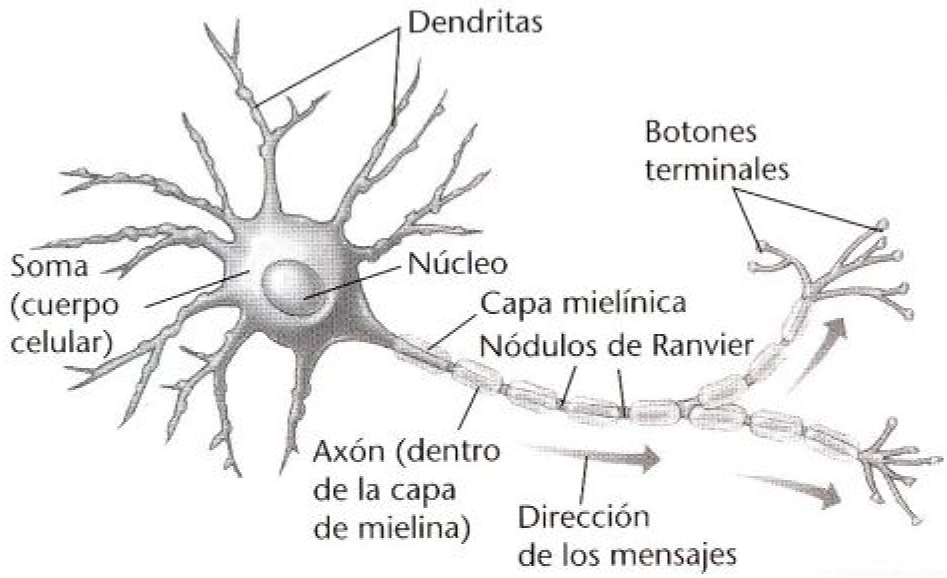
La teoría básica del perceptrón se presenta a continuación:

En un inicio las redes neuronales, tuvieron su inicio con las neuronas biológicas y se basaron en este elemento de la naturaleza para replicar parte de su funcionamiento, especialmente la identificación de patrones, por ejemplo en la siguiente imagen podemos ver una neurona biológica

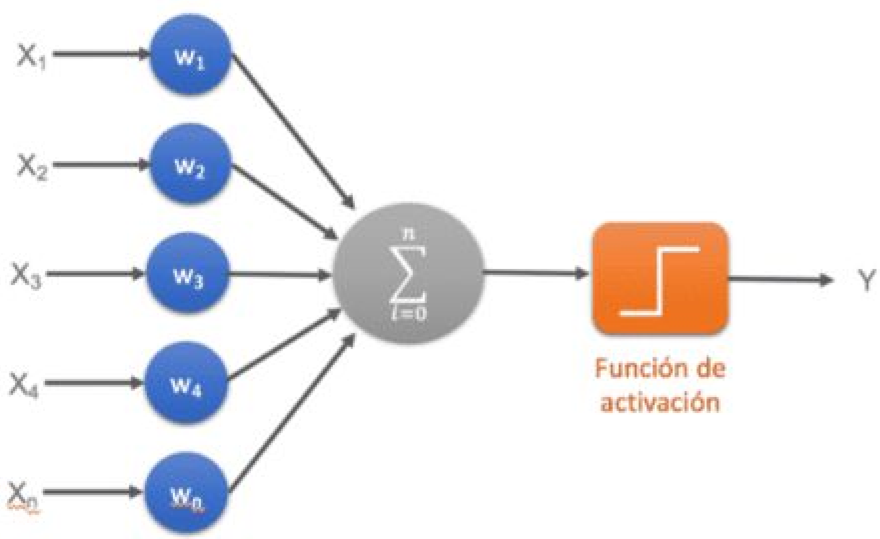


En la cual se ven tres partes que son las Dendritas, el axón y el soma que en nuestro lenguaje se representarán como las entradas, la salida y la función de sumatoria.

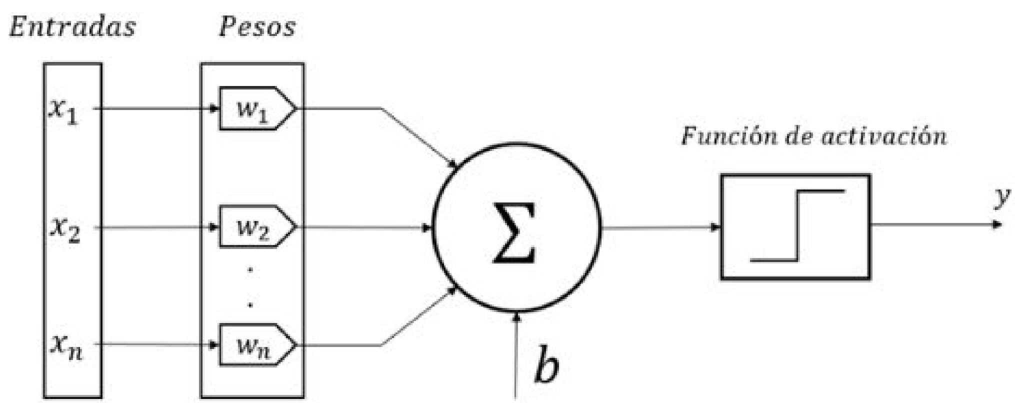
Aunque una Neurona biológica tiene más partes como se ve en la figura



Nosotros usaremos los tres elementos nombrados anteriormente más otros dos para poder hacer una identificación de patrones efectiva. los cuales serían las funciones de activación y los pesos



Además hay otro elemento importante que debemos tener en cuenta que es el perceptrón, el cual podemos definir cómo la forma más simple de una red neuronal usada para el reconocimiento de patrones (no muy complejos), siendo estos en su forma básica de la siguiente manera:



entonces podríamos describir al perceptrón como un dispositivo de computación con un umbral, que tiene en su estructura: una cantidad X_n de entradas y una cantidad W_n de pesos, un umbral, una función de activación y la salida.

Siendo el funcionamiento de aprendizaje del perceptrón un algoritmo iterativo que funciona de la siguiente forma:

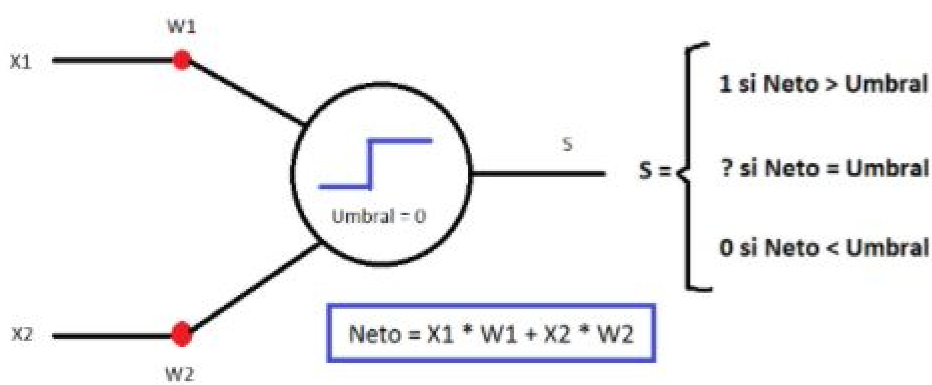
1. Se debe inicializar el valor del umbral y los pesos (En algunos casos podríamos iniciar los pesos de forma aleatoria)
2. Leer los valores de entrada
3. Hacer la sumatoria neta que es la multiplicación del valor de entrada con el peso asignado
4. pasar el valor de la sumatoria neta por el umbral, en este caso con un perceptrón simple, hacemos lo siguiente: si nuestro valor neto es por debajo del umbral su valor de salida es 0 si por el contrario el valor da por encima del umbral el valor será 1

Acá podemos ver un ejemplo de lo anterior dicho con la compuerta AND, en la cual vemos que cuando X1 y X2 son 1 se cumple el valor por encima del umbral, por consiguiente el valor de salida es 1

UMBRAL = U = 0.5

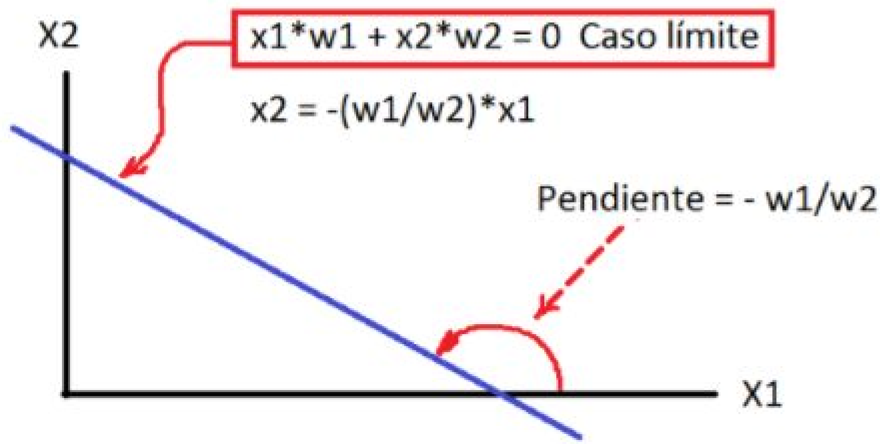
X1	X2	W1	W2	X1*W1+X2*W2 > U	Salida
0	0	0.3	0.3	0*0.3 + 0*0.3 = 0.0 NO	0
0	1	0.3	0.3	0*0.3 + 1*0.3 = 0.3 NO	0
1	0	0.3	0.3	1*0.3 + 0*0.3 = 0.3 NO	0
1	1	0.3	0.3	1*0.3 + 1*0.3 = 0.6 SI	1

Aunque siguiendo la teoría pudimos notas que: se aprecian sólo dos de los tres casos posibles a la hora de verificar el umbral, es decir, vimos que hacer cuando tenemos un valor neto mayor o menor al umbral, dejando por fuera el caso de que este fuera igual



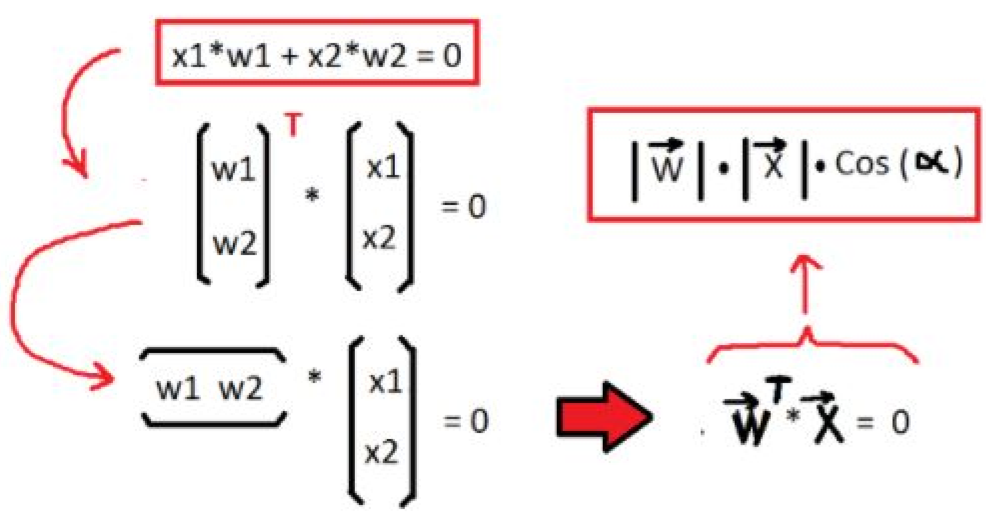
Entonces tratemos de entender este punto en donde el valor neto es igual al umbral, nosotros conocemos que el valor neto es la sumatoria de multiplicar el valor de la entrada por sus correspondientes pesos, siendo X1 y X2 valores estáticos y W1 y W2 nuestros valores variables, al ser los pesos que vamos a iterar.

entonces,¿cómo podemos mostrar matemáticamente una neurona ? la respuesta es sencilla y también es fácil de ver.



es a través de una recta, siendo x2 = -(w1/w2) * x1 la función de esta representación y la pendiente = -w1 / w2 que es cambiante de acuerdo a los valores que tengan los pesos

Como información adicional y necesaria para entender el funcionamiento de las redes neuronales, debemos saber que los valores de entrada y los pesos no son sólo una magnitud, es decir, estos son vectores y la operación que hacemos entre ellos es una una multiplicación de vectores.Pero para hacer la multiplicación de forma correcta primero debes pasar el vector de pesos a su transpuesta y daría lo siguiente



Por los tanto el producto punto de esos dos vectores es igual al producto punto (producto de sus magnitudes) por el coseno del ángulo, teniendo que la recta solución es aquella que el producto punto es igual a cero, es decir cuando el ángulo es 90°.

$$\boxed{|\vec{w}| \cdot |\vec{x}| \cdot \mathrm{Cos}(\propto)} = 0$$

4 LÓGICA DIFUSA - INTRODUCCIÓN

Lotfi Zadeh, azerbaiyano que adelantó estudios en la Universidad de Teherán y en el MIT además de las universidades de Columbia y Berkeley, donde también ejerció, investigó sobre lo que significó el primer paso para la investigación sobre la lógica difusa, el principio de incompatibilidad, esto a mediados de los años sesenta en la Universidad de Berkeley. Este principio fue definido como: "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".

Con ello trajo consigo el concepto sobre los conjuntos difusos "Fuzzy Set", este considera que el pensamiento humano se construye mediante etiquetas lingüísticas en lugar de números. Esta lógica sin embargo cuenta con el trabajo en conjunto de datos numéricos y términos lingüísticos, permite la representación del conocimiento común, que cuenta con características lingüísticas cualitativas y no obligatoriamente cuantitativas.

La diferencia con los sistemas basados en la lógica clásica es que estos pueden reproducir de manera admisible la manera usual en que se razona teniendo en cuenta que la certeza de una proposición depende de un grado, haciendo referencia a los principios formales de razonamiento aproximado. La flexibilidad, tolerancia con la imprecisión. la capacidad de modelamiento de problemas no lineales y su base en lenguaje natural son algunas de sus características.

Los orígenes de la lógica difusa datan de 2500 años atrás a 1965, Aristóteles precedió con la consideración de la existencia de los grados de veracidad y falsedad, Platón hizo lo propio usando grados de pertenencia.

Este tipo de lógica supuso una resistencia en la comunidad científica, sin embargo muchos investigadores siguieron de cerca las teorías de Zadeh, lo que trajo consigo el establecimiento de algunos grupos de investigación en varias universidades de Japón, logrando grandes contribuciones al desarrollo de la teoría y el estudio de sus aplicaciones.

En 1974 se registró un acontecimiento importante para la lógica difusa en Reino Unido, Asillian y Mamdani desarrollaron un controlador difuso para una máquina de vapor, sin embargo, en 1980 se implantó por primera vez en una planta cementera, F.L. Smidth & Co. fueron los responsables. Siguiendo en la misma línea Fuji usa este tipo de lógica para el control de inyección química en unas plantas depuradoras de agua, en el año de 1983. Finalmente en el año 1987 Hitachi usa un controlador difuso para el control del metro de Sendai.

En los años 80 los investigadores se aproximan para la creación de reglas fuzzy a partir de datos de entrenamiento. Además, perciben la similitud entre las redes neuronales y los sistemas fuzzy, el interés en las primeras estaba creciendo. Iniciaron buscando similitudes entre ambas técnicas, los resultados de estas son denominados neuro-fuzzy systems. Kosko es un gran contribuyente a esta área.

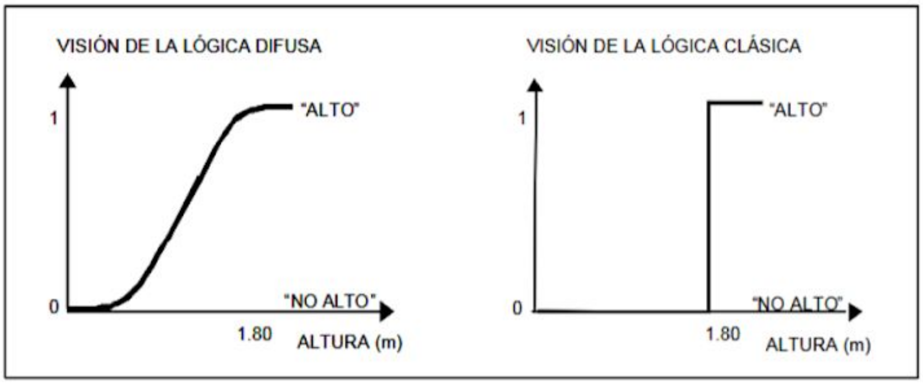
En la siguiente década aparecen los algoritmos genéticos, uniendolos a las técnicas mencionadas anteriormente, se conforman como herramientas para los sistemas de control.

No obstante, Zadeh pretendía contribuir en la creación de un formalismo para tratar la imprecisión que trae consigo el pensamiento humano expresado en términos lingüísticos, siendo una sorpresa el enfoque que tuvo en Japón sobre el control automático de procesos, logrado por la cooperación entre el gobierno, las universidades e industrias japonesas durante varios años.

A través de los años han sido lanzados muchos productos al mercado con el uso de esta tecnología, la etiqueta fuzzy se convirtió en un símbolo de calidad. La metalurgia, los robots de fabricación, controles de maniobras de aviones, sistemas para la estabilización de imagen de algunas empresas (Sony, Sanyo y Cannon), lavadoras autorreguladoras de la cantidad de jabón necesaria para un lavado teniendo en cuenta el grado de suciedad de la ropa (Panasonic y Bosch), control de cambio automático para frenar en casos críticos de Renault, lavaplatos, sistema para regular la cantidad de anestesia suministrada a un paciente, sistemas de entrega o denegación de créditos, entre otros.

Además, hay muchas áreas en la investigación de la lógica difusa, entre ellas el reconocimiento de patrones visuales o la identificación de segmentos de ADN. El uso de esta tecnología podría considerarse el futuro para el tratamiento de datos, recuperación de información y control de la red.

Zadeh ilustra con el conjunto de los hombres altos, en la lógica clásica se establece que este conjunto podría estar conformado por hombres con estatura mayor a 1.80 metros, una persona que mida 1.79 metros estaría fuera de este conjunto aunque esté apenas a dos centímetros de alguien que mide 1.81 metros, se consideraba esto algo poco lógico. Por ello la lógica difusa establece que no existe una clara diferencia para la pertenencia o no pertenencia a este conjunto, así que entre valores de 0 y 1, hombres de distintas alturas pueden tener cierto grado de pertenencia al grupo de hombres altos. En el siguiente gráfico se pueden apreciar ambas visiones:



El grado de pertenencia a un conjunto difuso puede tomar cualquier valor entre 0 y 1, este se define a través de la función característica que esté asociada al conjunto. Para cada valor que pueda tomar la variable de entrada x, la función característica da el grado de pertenencia del valor de x al conjunto difuso A.

Un conjunto clásico se puede definir enumerando los elementos que pertenecen al conjunto, clarificando las propiedades que deben cumplir los elementos pertenecientes al conjunto o en términos de la función de pertenencia.

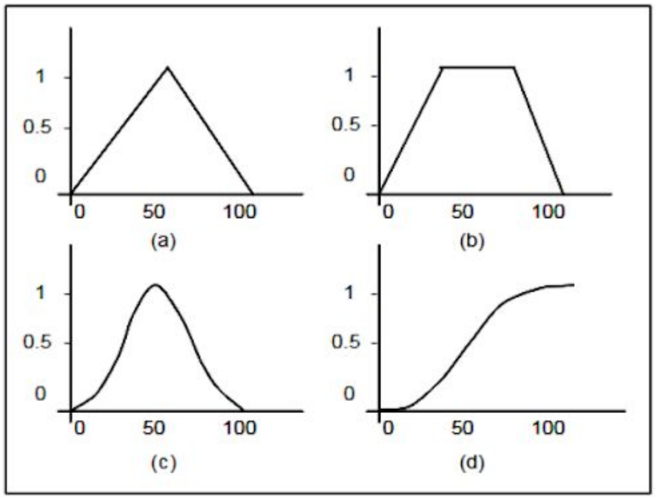
$$\propto_A(x) = \begin{Bmatrix} 1 \text{ si } \mathrm{x} \in A \\ 0 \; si \; \mathrm{x} \notin \mathrm{A} \end{Bmatrix}$$

Un conjunto difuso en un universo de discurso U se puede representar como un conjunto de pares ordenados de un elemento 'x' junto a su valor de pertenencia al conjunto.

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

La función característica da una medida del grado de coincidencia de un elemento de U con el conjunto difuso, la forma de esta depende del criterio que se aplique a la solución de un problema y en variación a la interpretación del usuario. Debe cumplir la condición de tomar valores entre 0 y 1 con continuidad. Entre estas funciones se encuentran la triangular, trapezoidal, gaussiana, sigmoidal entre otras.

Para determinar la función característica asociada a un conjunto, hay dos aproximaciones, aquella basada en el conocimiento humano y la otra basada en la utilización de una colección de datos. Algunos ejemplos son:


1
0.5
0
0
50
100
(a)
1
0.5
0
0
50
100
(b)
1
0.5
0
0
50
100
(c)
1
0.5
0
0
50
100
(d)

5 CONCLUSIONES

El desarrollo de las temáticas elaboradas en clase utilizando el lenguaje Python prueba ser un mecanismo de gran valor para el aprendizaje de los conceptos básicos de la materia al igual que la realización de monografías que permiten un mejor entendimiento.

6 BIBLIOGRAFÍA

https://repl.it

http://bibing.us.es/proyectos/abreproy/11084/fichero/Memoria+por+cap%C3%ADtulos+%252FCap%C3%ADtulo+4.pdf+

https://www.diegocalvo.es/perceptron/