

Computación Blanda

Soft Computing

Autor: Julian Giraldo Cardona

IS&C, Universidad Tecnológica de Pereira, Pereira, Colombia

Correo-e: julian.giraldo2@utp.edu.co

Resumen— Este documento presenta un resumen de las líneas clásicas de la Computación Blanda: redes neuronales, lógica difusa, sistemas expertos, algoritmos genéticos y machine learning. El objetivo del documento es brindar una panorámica general de las temáticas, mostrando su relación con las técnicas de inteligencia artificial. La diferencia entre el paradigma de Inteligencia Artificial y la computación blanda está centrada en el mecanismo de inferencia utilizado y su aplicación a la solución de problemas tomados de lo cotidiano, de las teorías de conocimiento y de su relación con ciencias afines.

Palabras clave— sistemas, redes, inteligencia artificial, software, computación, investigación, industria, genético, aprendizaje.

Abstract— This document presents a summary of the classic lines of Soft Computing: neural networks, fuzzy logic, expert systems, genetic algorithms and machine learning. The objective of the document is to provide a general overview of the topics, showing their relationship with artificial intelligence techniques. The difference between the Artificial Intelligence paradigm and soft computing is centered on the inference mechanism used and its application to the solution of problems taken from everyday life, from knowledge theories and their relationship with related sciences.

Key Word— systems, networks, artificial intelligence, software, computing, research, industry, genetic, learning.

I. INTRODUCCIÓN

La temática de la Computación Blanda se encuentra enmarcada en el paradigma de la Inteligencia Artificial. La diferencia con dicho paradigma radica en que la Computación Blanda está centrada en la aplicación pragmática de las teorías de la Inteligencia Artificial a la solución de problemas complejos en diversos campos del conocimiento.

Las líneas derivadas de la Computación Blanda, se configuran en las siguientes tendencias: a) Redes Neuronales Artificiales, b) Lógica Difusa, c) Sistemas Expertos, d) Algoritmos Genéticos, e) Deep Learning (Machine Learning).

En los siguientes apartados se presenta un resumen de dichas tendencias.

1.1 REDES NEURONALES

Las redes neuronales se han convertido en la familia de algoritmos machine learning más populares, como modelo computacional, las redes neuronales existen desde mediados del siglo pasado pero ahora último con la mejora de la tecnología hemos explotado su potencial al máximo.

Reconocimiento de caracteres, de imágenes, de voz, predicción versátil, generación de texto, traducción de idiomas, conducción autónoma, análisis genético.

La complejidad de estas redes neuronales surge de unas partes más pequeñas trabajando en conjunto formando así un sistema. Cómo es en este caso a cada parte de este conjunto de la red se le llama neurona.

Similar a una neurona biológica, estas neuronas son capaces de recibir información por un extremo llamados también valores de entrada y al otro extremo tiene un valor de salida. Algo así como una función matemática.

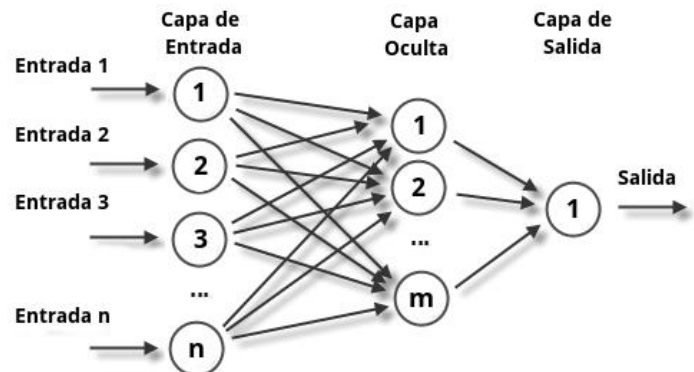


Fig.1 Red Neuronal

Como se mencionaba anteriormente y como lo muestra la Fig.1 Una red neuronal consta de muchas neuronas conectadas entre sí para entregar un resultado final. Esto puede ser representado en la parte de las heurísticas para llegar a un resultado preciso. Cada neurona hace su proceso con las entradas que recibe e internamente hace una suma ponderada de cada una y este proceso viene asignado con el peso que se le asigna a cada una de las entradas que se le asigna, este peso indica la intensidad de cada variable que recibe la neurona.

El objetivo principal de este modelo es aprender modificándose automáticamente a sí mismo de forma que puede llegar a realizar tareas complejas que no podrían ser realizadas mediante la clásica programación basada en reglas. De esta forma se pueden automatizar funciones que en un principio solo podrían ser realizadas por personas.

Existe un simple ejemplo de nuestra vida cotidiana. Si me siento débil y tengo vomito entonces estoy enfermo, se hace un modelo de variables binarias 1 o 0.

Me siento débil: X_1 1 | 0
Tengo vomito: X_2 1 | 0

Donde el valor binario “1” representa que sí estoy débil y que sí tengo vomito. Por otro lado también se espera una variable de salida Y_1 que también cuenta con los valores binarios para afirmar si estamos enfermos o no.

X_1	X_2	Resultado
0	0	0
1	0	0
0	1	0
1	1	1

Fig.2 Tabla operación neuronal

Funciona completamente igual a una compuerta lógica AND con una serie de parámetros con sus valores binarios como se observa en la Fig.2

Existen diferentes capas en una red neuronal (Fig.1) Todo depende en la posición en que esté ubicada cada neurona ya que una puede recibir el valor procesado por la anterior, con este proceso la red puede aprender conocimiento jerarquizado, entre más capas añadimos más complejo puede ser el sistema.

Algunas de las características de las redes neuronales son:

Redes Monocapa: Se establecen conexiones laterales, cruzadas o autor recurrentes entre las neuronas que pertenecen a la única capa que constituye la red. Se utilizan en tareas relacionadas con lo que se conoce como auto asociación; por ejemplo, para generar informaciones de entrada que se presentan a la red incompleta o distorsionada.

Redes Multicapa: Son aquellas que disponen de conjuntos de neuronas agrupadas en varios niveles o capas. Una forma de distinguir la capa a la cual pertenece la neurona, consiste en fijarse en el origen de las señales que recibe a la entrada y el destino de la señal de salida. Según el tipo de conexión, como se vio previamente, se distinguen las redes feedforward, y las redes feedforward/feedback.

Topología de las redes neuronales: la estructura de las redes neuronales consiste en una organización de las neuronas formando capas alejadas y no tan alejadas de la entrada y la salida de la red, en este orden de ideas los parámetros de la red son el número de capas, el número de neuronas por capa, el grado de conectividad y el tipo de conexiones entre neuronas.

Mecanismo de aprendizaje: La red neuronal tiene la capacidad de modificar sus pesos en respuesta a una información de entrada, los cambios que ocurren en el proceso de aprendizaje significan una destrucción, modificación y creación de conexiones entre las neuronas. La creación de una nueva conexión implica que el peso de la misma pasa a tener un valor distinto de cero, una conexión se destruye cuando su peso pasa a ser cero.

Ventajas de las redes neuronales: Debido a su tipología, las redes neuronales presentan un gran parecido con el cerebro humano y no solo estructuralmente sino que también a la hora de aprender algo, de obtener experiencia para corregir errores, de abstraer características esenciales a partir de entradas que representan información irrelevante. Entre estas ventajas están:

Aprendizaje: Una red neuronal tiene la habilidad de aprender mediante entradas que uno le transmite indicando que resultado en la salida se espera.

Auto Organización: Una red neuronal crea su propia representación de la información en su interior.

Tolerancia a fallos: Gracias a que la red neuronal funciona de una forma redundante, ésta puede seguir respondiendo correctamente aun si tiene algún daño.

Flexibilidad: La red neuronal tiene una gran flexibilidad a la hora de manejar los valores de entrada porque aun así los procesa con ruido y todo.

Tiempo real: Gracias a que la red neuronal trabaja de forma paralela entonces es fácil la comunicación entre diferentes dispositivos y pueden trabajar en tiempo real.

Aplicaciones en redes neuronales:

Se pueden desarrollar redes neuronales en un periodo de tiempo razonable, con la capacidad de realizar tareas concretas mejor que otras tecnologías. Cuando se implementan mediante hardware (redes neuronales en chips VLSI), presentan una alta tolerancia a fallos del sistema y proporcionan un alto grado de paralelismo en el procesamiento de datos. Esto posibilita la inserción de redes neuronales de bajo coste en sistemas existentes y recientemente desarrollados.

Hay muchos tipos diferentes de redes neuronales; cada uno de los cuales tiene una aplicación particular más apropiada. Algunas aplicaciones comerciales son:

Biología: Aprender más acerca del cerebro y otros sistemas.

Empresa: Evaluación de probabilidad de formaciones geológicas y petrolíferas, explotación de bases de datos.

Medioambiente: Analizar tendencias y patrones.

Finanzas: Previsión de la evolución de los precios, identificación de falsificaciones.

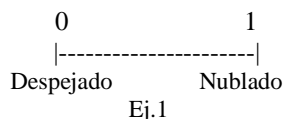
Manufacturación: Robots automatizados y sistemas de control (visión artificial y sensores de presión, temperatura, gas, etc.).

Medicina: Analizadores del habla para ayudar en la audición de sordos profundos.

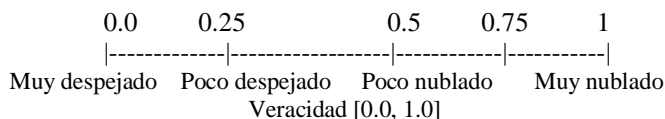
Militares: Clasificación de las señales de radar, reconocimiento y seguimiento en el tiro al blanco.

1.2 LÓGICA DIFUSA

En la lógica clásica, está fundamentada en la lógica aristotélica donde solo existen dos valores de veracidad (1,0) de esta forma tomamos decisiones en la vida cotidiana. La lógica difusa propone darle una gama de posibilidades de veracidad teniendo más de dos estados, puede existir que una afirmación sea parcialmente verdadera y parcialmente falsa, puede tomar valores entre 0 y 1. Por ejemplo: ¿Qué tan nublado está el cielo? En la lógica clásica sólo existe la posibilidad de que el cielo está nublado o no está nublado.



En la lógica difusa este ejemplo puede tener hasta 4 estados diferentes.



Se nota una gran diferencia entre lógica clásica y lógica difusa por su amplio rango de veracidad. La lógica difusa es útil para automatizar procesos y controladores difusos, así mismo se puede trasladar la experiencia hacia una computadora. La experiencia humana como la de un ingeniero o un trabajador que ha estado muchos años familiarizado con máquinas puede enseñarle mediante la experiencia a un ordenador por medio de una serie de reglas en forma lingüística para que así lo pueda entender la máquina y así se crea un control difuso.

Conjuntos difusos y funciones Características:

Uno de los primeros ejemplos utilizados por Lofti A. Zadeh, para ilustrar el concepto de conjunto difuso, fue el conjunto “hombres altos”. Según dice la teoría de la lógica clásica el conjunto de “hombres altos” es un conjunto al que pertenecerían los hombres con la estatura mayor a un cierto umbral, que podríamos decir que es 1.80 metros, por ejemplo. Entonces todos los hombres con una altura inferior a este valor quedarían fuera del conjunto de “hombres altos”. Así tendríamos que un hombre mide 1.81 metros de estatura pertenecería al conjunto de hombres altos, y en cambio un hombre que mida 1.79 ya no pertenecería a este grupo como se mencionaba anteriormente en el Ej.1

Sin embargo, no parece muy lógico decir que un hombre es alto y otro no lo es cuando la diferencia de altura es solamente por 2 cm. 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 0.85 y uno que mida 1.60m con un grado 0.1. Analizando de esa forma se

puede considerar que la lógica clásica es un caso límite de la lógica difusa en el que se asigna un grado de pertenencia 1 a los hombres con altura mayor o igual a 1.80 y un grado de pertenencia 0 a los que tienen una altura menor. (Fig. 3)

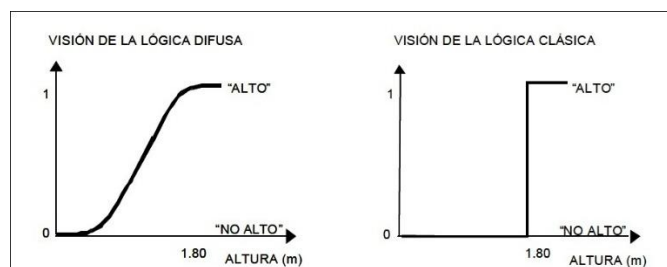


Fig. 3 Lógica clásica vs lógica difusa

En este orden de ideas, los conjuntos difusos pueden ser considerados como una generalización de los conjuntos clásicos. 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. Este grado de pertenencia o veracidad se define mediante la función característica asociada al conjunto difuso: para cada valor que pueda tomar un elemento o variable de entrada x la función característica $\mu_A(x)$ proporciona el grado de pertenencia de ese valor de x al conjunto difuso A .

Ventajas de la lógica difusa:

- La principal ventaja de este sistema de control es su facilidad de implementación.
- Este tipo de sistemas están dando muy buenos resultados en procesos no lineales y de difícil modelación.
- El modo de funcionamiento es similar al comportamiento humano.
- Es una forma rápida y económica de resolver un problema.
- No se necesita conocer el modelo matemático que rige su funcionamiento.
- Puede tomar decisiones que a nosotros nos cuesta mucho procesarlas.

Aplicaciones de lógica difusa:

- Videocámaras y cámaras fotográficas.
- Control de sistemas de vehículos (frenos, embragues, aire acondicionado).
- Control de procesos biológicos, químicos, económicos.
- Cargadores de baterías, puentes grúa.
- Lavadoras, secadoras.

1.3 SISTEMAS EXPERTOS

Los sistemas expertos se han utilizado desde hace mucho tiempo pero en los últimos años es cierto que se ha dejado a un lado este método pero aun así sigue siendo fundamental para la toma de decisiones. Las características de un sistema experto son:

- Toma decisiones con la competencia de un experto
- Representa el conocimiento y realiza inferencias
- Basado en gran volumen de conocimiento teórico
- Razona sobre las operaciones que realiza, cómo y porqué las ejecuta
- Constituye un modelo no algorítmico
- Admite una fácil modificación del conocimiento sin reprogramar el sistema
- El usuario dialoga con el sistema empleando un lenguaje natural limitado

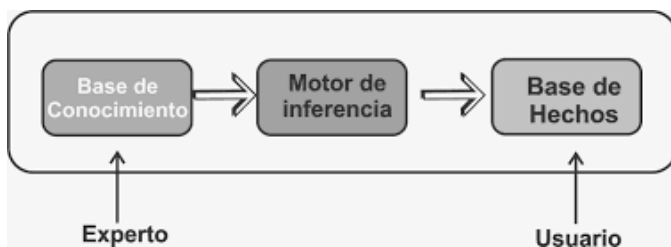


Fig.4 Estructura de sistema experto

El motor de inferencia lo que va a hacer es recorrer la base de datos de conocimiento para intentar dar una solución a una pregunta que hagamos desde la interfaz de usuario, la base de conocimiento almacena todo el conjunto de reglas.

Sistema experto con encadenamiento hacia adelante:

Un grupo de múltiples reglas de inferencia que contiene un problema con su solución se llama cadena.

Una cadena que realiza una búsqueda o camino desde el problema a la solución se llama una cadena hacia adelante.

Este tipo de cadena va desde los hechos hasta las conclusiones que siguen a partir de los hechos.

Una cadena que transita hacia atrás desde una hipótesis hasta los hechos que soportan tal hipótesis se llama una cadena hacia atrás. Otra forma de definirla es en términos de una meta que puede estar formada por submetas que se han cumplido.

El encadenamiento se puede expresar con cierta facilidad en términos de inferencia si suponemos que tenemos reglas del tipo Modus-poner. En un ejemplo de encadenamiento hacia adelante tenemos:

Base inicial de hechos: H, K

Base de reglas: R1: $A \rightarrow E$
 R2: $B \rightarrow D$
 R3: $H \rightarrow A$
 R4: $E \text{ y } G \rightarrow C$
 R5: $E \text{ y } K \rightarrow B$
 R6: $D \text{ y } E \text{ y } K \rightarrow C$
 R7: $G \text{ y } K \text{ y } F \rightarrow A$

La pregunta sería, ¿Existe D? dados los hechos H, K, Entonces la idea sería ir recorriendo cada regla para inferir en el resultado.

Reglas:

R3: $H \rightarrow A$

R1: $A \rightarrow E$

R5: $E \text{ y } K \rightarrow B$

R2: $B \rightarrow D$ Respuesta favorable

R6: $D \text{ y } E \text{ y } K \rightarrow C$

En la regla 6 (R6) existe una saturación

Base de hechos

- A, H, K

- A, E, H, K

- A, B, E, H, K

- A, B, D, E, H, K

- A, B, C, D, E, H, K

En este encadenamiento hacia adelante se hace de una recursiva avanzando así en cada regla hasta encontrar el resultado esperado y llegamos a la conclusión que D sí es cierto, en este sistema se puede avanzar hasta llegar a una saturación.

Encadenamiento hacia atrás:

En el encadenamiento hacia atrás el proceso es el inverso: queremos probar la hipótesis de animal(Bugs-Bunny). El problema central del encadenamiento hacia atrás es encontrar una cadena de enlace entre la evidencia y la hipótesis.

El hecho conejo(Bugs Bunny) se llama la evidencia en el encadenamiento hacia atrás, para indicar que se usará para sostener la hipótesis, siendo esta evidencia usada para probar dicha hipótesis. Tenemos una estructura de cláusulas.

Los encadenamientos hacia adelante y hacia atrás son en realidad caminos a través de un espacio de estados de un problema. En este espacio de estados del problema los estados intermedios se corresponden con hipótesis intermedias bajo el encadenamiento hacia atrás o conclusiones intermedias bajo el encadenamiento hacia adelante.

Conservamos la base inicial de hechos y la base de reglas para desarrollar este ejemplo.

Conejo (Bugs Bunny)

$\text{Conejo}(x) \rightarrow \text{Mamífero}(x)$

$\text{Mamífero}(x) \rightarrow \text{Animal}(x)$

Animal (Bugs Bunny)

Conservamos la base inicial de hechos y la base de reglas para desarrollar este ejemplo.

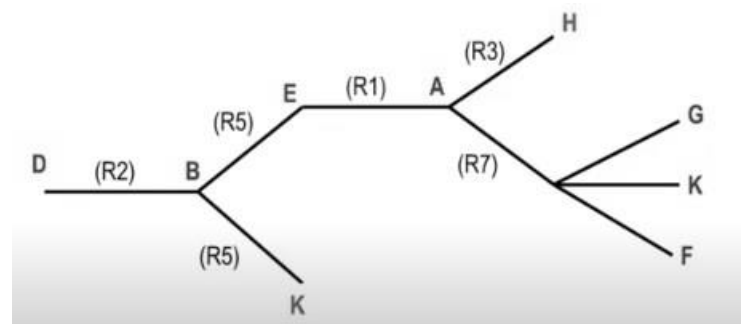


Fig.5 Árbol encadenamiento hacia atrás

1.4 ALGORITMOS GENÉTICOS

Los Algoritmos Genéticos (AGs) son métodos adaptativos que pueden usarse para resolver problemas de búsqueda y optimización. Están basados en el proceso genético de los organismos vivos. A lo largo de las generaciones, las poblaciones evolucionan en la naturaleza de acuerdo con los principios de la selección natural y la supervivencia de los más fuertes, postulados por Darwin (1859).

Por imitación de este proceso, los Algoritmos Genéticos son capaces de ir creando soluciones para problemas del mundo real. La evolución de dichas soluciones hacia valores óptimos del problema depende en buena medida de una adecuada codificación de las mismas.

Los principios básicos de los Algoritmos Genéticos fueron establecidos por Holland (1975), y se encuentran bien descritos en varios textos – Goldberg (1989), Davis (1991), Michalewicz (1992),

Reeves (1993) – .En la naturaleza los individuos de una población compiten entre sí en la búsqueda de recursos tales como comida, agua y refugio. Incluso los miembros de una misma especie compiten a menudo en la búsqueda de un compañero. Aquellos individuos que tienen más éxito en sobrevivir y en atraer compañeros tienen mayor probabilidad de generar un gran número de descendientes. Por el contrario individuos poco dotados producirán un menor número de descendientes. Esto significa que los genes de los individuos mejor adaptados se propagarán en sucesivas generaciones hacia un número de individuos creciente. La combinación de buenas características provenientes de diferentes ancestros, puede a veces producir descendientes “súper individuos”, cuya adaptación es mucho mayor que la de cualquiera de sus ancestros. De esta manera, las especies evolucionan logrando unas características cada vez mejor adaptadas al entorno en el que viven.

Los Algoritmos Genéticos usan una analogía directa con el comportamiento natural. Trabajan con una población de individuos, cada uno de los cuales representa una solución factible a un problema dado. A cada individuo se le asigna un valor o puntuación, relacionado con la bondad de dicha solución. En la naturaleza esto equivaldría al grado de efectividad de un organismo para competir por unos determinados recursos. Cuanto mayor sea la adaptación de un individuo al problema, mayor será la probabilidad de que el mismo sea seleccionado para reproducirse, cruzando su material genético con otro individuo seleccionado de igual forma. Este cruce producirá nuevos individuos – descendientes de los anteriores – los cuales comparten algunas de las características de sus padres.

Cuanto menor sea la adaptación de un individuo, menor será la probabilidad de que dicho individuo sea seleccionado para la reproducción, y por tanto de que su material genético se propague en sucesivas generaciones.

De esta manera se produce una nueva población de posibles soluciones, la cual reemplaza a la anterior y verifica la interesante propiedad de que contiene una mayor proporción de buenas características en comparación con la población anterior. Así a lo largo de las generaciones las buenas características se propagan a través de la población. Favoreciendo el cruce de los individuos mejor adaptados, van siendo exploradas las áreas más prometedoras del espacio de búsqueda. Si el Algoritmo Genético ha sido bien diseñado, la población convergerá hacia una solución óptima del problema.

Proceso de selección + cruce:

En el proceso de selección se hace mediante una ruleta de forma aleatoria eligiendo cada padre y posteriormente elegir los que tengan los genes más fuertes de toda la población.



Fig.6 Ruleta con porcentajes para selección

Se crea un esquema (scheme) con los genes más fuertes que contienen los cromosomas más importantes de la población. Por lo general son 3 valores en el esquema pero pueden ser 4 o más, dependiendo del análisis que se quiera realizar.

Durante la fase reproductiva se seleccionan los individuos de la población para cruzarse y producir descendientes, que constituirán, una vez mutados, la siguiente generación de individuos. La selección de padres se efectúa al azar usando un procedimiento que favorezca a los individuos mejor adaptados, ya que a cada individuo se le asigna una probabilidad de ser seleccionado que es proporcional a su función de adaptación. Este procedimiento se dice que está basado en la ruleta sesgada. Según dicho esquema, los individuos bien adaptados se escogerán probablemente varias veces por generación, mientras que los pobremente adaptados al problema, no se escogerán más que de vez en cuando.

Una vez seleccionados dos padres, sus cromosomas se combinan, utilizando habitualmente los operadores de cruce y mutación. Las formas básicas de dichos operadores se describen a continuación.

El operador de cruce, coge dos padres seleccionados y corta sus ristas de cromosomas en una posición escogida al azar, para producir dos subristas iniciales y dos subristas finales. Después se intercambian las subristas finales, produciéndose dos nuevos cromosomas completos.

Ambos descendientes heredan genes de cada uno de los padres. Este operador se conoce como operador de cruce basado en un punto. Habitualmente el operador de cruce no se aplica a todos.

Padres 1 0 1 0 | 0 0 1 1 1 0 0 0 1 1 | 0 1 0 0 1 0

Hijos 1 0 1 0 0 1 0 0 1 0 0 0 1 1 0 0 1 1 1 0

El punto de cruce se establece por “|” con dos padres seleccionados se generan 2 hijos con sus genes intercambiados para continuar la descendencia.

La mutación en algoritmos genéticos:

Se supone que los individuos (posibles soluciones del problema), pueden representarse como un conjunto de parámetros (que denominaremos genes), los cuales agrupados forman una ristra de valores (a menudo referida como cromosoma). Si bien el alfabeto utilizado para representar los individuos no debe necesariamente estar constituido por el {0, 1}, buena parte de la teoría en la que se fundamentan los Algoritmos Genéticos utiliza dicho alfabeto. En la mutación se tiene un índice que puede variar entre 0 y 100% y se aplica en una población. La selección para saber que individuo mutar se puede realizar mediante forma aleatoria o también haciendo un control dependiendo en que puede favorecer dicha mutación. Por ejemplo tenemos 1 cromosoma con 9 diferentes bivalores.

1 0 0 1 0 1 0 1 1 1
1 1 0 1 0 1 0 1 1 1

En la mutación simplemente se cambia el valor de 0 a 1.

Supervivencia de los esquemas:

Después de tener los esquemas seleccionados se hacen operaciones estadísticas incluyendo la longitud del cromosoma, el número de individuos seleccionados en toda la población, el orden $O(s)$, la longitud $\&(s)$, el promedio de supervivencia de los esquemas seleccionados, el eval promedio, y el porcentaje de selección. Cada uno de los procesos en que se enfrenta la población se hace las respectivas operaciones para así determinar qué tan posible es que sobrevivan. Por lo general, si el procedimiento se hace bien, este produce una gráfica exponencial.

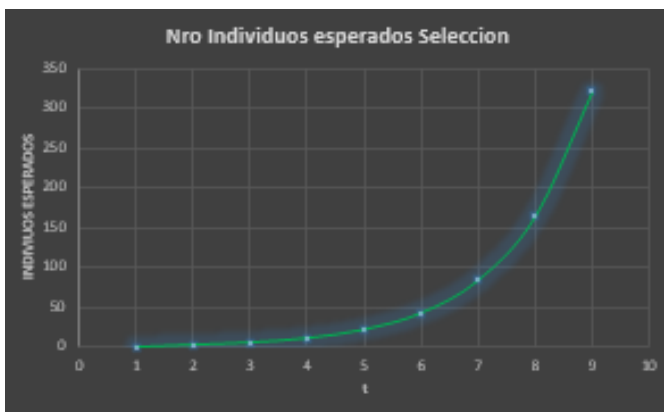


Fig.7 gráfico de supervivencia de los esquemas Alg Genéticos

1.5 DEEP LEARNING

El Deep learning se basa en redes neuronales, las redes neuronales empezaron en los años 60s y 70s, recién fue en los años 80s que se hicieron famosas y luego se detuvo su avance ya que en ese entonces no había los procesadores necesarios y tampoco contábamos con la cantidad de datos. El Deep learning intenta replicar el cerebro humano por medio de las redes neuronales, el mayor problema es que nuestro cerebro tiene millones y millones de neuronas.

En la Fig.1 se aprecian los nodos, estos nodos hacen el trabajo igual que las neuronas como se explica en redes neuronales (1.1).

Diferencias entre Inteligencia Artificial, Machine Learning y Deep Learning:

La Inteligencia Artificial (IA) es un subcampo de la informática que se creó en la década de 1960, y que trata de solucionar tareas que son sencillas para los seres humanos, pero difíciles para las computadoras. Se trata de un concepto bastante genérico e incluye todo tipo de tareas tales como la planificación, el reconocimiento de objetos y sonidos, hablar, traducir, realizar actividades creativas (como por ejemplo crear obras de arte, o la poesía), etc.

El Machine Learning o aprendizaje automático se ocupa de un aspecto de la IA: dado un problema de IA que se puede describir en términos discretos (por ejemplo, dado un conjunto de posibles acciones para un determinado objetivo, cuál es la correcta), y disponiéndose de una gran cantidad de información sobre el mundo, se determina cual es la acción "correcta", sin que el mecanismo de elección se encuentre previamente programado. Es decir, el sistema aprende de forma autónoma a tomar las decisiones. De forma práctica esto se traduce en una función en la que a partir de una entrada se obtiene una salida, con lo que el problema radica en construir un modelo de esta función matemática de forma automática. Por lo tanto, la principal diferencia radica en que un programa muy inteligente que tiene un comportamiento similar al humano puede ser IA, pero a menos que sus parámetros se aprendan automáticamente a partir de los datos, no es Machine Learning.

Machine Learning:

Aunque los términos se utilizan a veces como sinónimos, el Deep Learning y el Machine Learning no son lo mismo, siendo el primero un tipo particular del segundo, es decir, el Deep Learning es Machine Learning, pero existen técnicas de Machine Learning que no son Deep Learning.

El Machine Learning, como se ha visto, se describe a menudo como un tipo de técnicas de Inteligencia Artificial donde las computadoras aprenden a hacer algo sin ser programadas para ello. Por poner un ejemplo sencillo, se podría programar un ordenador para identificar a un animal como un gato escribiendo un código que indique al programa que elija "gato" cuando se ve una imagen concreta de un gato. Esto funcionaría si el único gato con el que tratase el programa es el de esa imagen, pero no lo haría si el programa tuviera que ver un

montón de imágenes de diferentes animales, incluyendo una gran cantidad de gatos, y tuviera que identificar cuáles de ellas representan a un gato.

Los programas de Machine Learning actuarían en ese segundo caso, para lo cual pueden entrenarse de diferentes maneras. En una de ellas al programa se le muestra una gran cantidad de imágenes de diferentes animales (etiquetadas con el nombre del animal correspondiente). El programa aprenderá que los animales que se parecen a los gatos se llaman "gato" sin haber sido programado para llamar "gato" a una imagen concreta de un gato. Para ello el programa debe aprender combinaciones de características visuales que tienden a aparecer juntas (por ejemplo, la forma de los cuerpos y de las caras), haciendo a los gatos visualmente diferentes de otros animales. El programa aprende a asociar esta combinación de características distintivas con la palabra "gato", proceso de aprendizaje conocido como "construcción de un modelo de un gato".

Una vez que se ha construido el modelo de "gato", un programa de Machine Learning prueba dicho modelo al tratar de identificar a los gatos en un conjunto de imágenes que no ha visto antes. El programa mide el éxito obtenido en la identificación de los nuevos gatos y utiliza esta información para ajustar el modelo, de forma que la próxima vez obtendrá mejores resultados. Es decir, el nuevo modelo se prueba, se evalúa su rendimiento, y se realiza otro ajuste. Este proceso iterativo continúa hasta que el programa haya construido un modelo que pueda identificar gatos con un alto nivel de precisión.

Si el entrenamiento es "supervisado", en cada paso iterativo de prueba y perfeccionamiento del modelo se compara la etiqueta asignada a la foto con la "decisión" tomada por el programa, determinando si ha identificado el tipo de imagen correctamente. El entrenamiento supervisado es relativamente rápido y necesita relativamente pocos recursos computacionales. Sin embargo necesita de la intervención humana para etiquetar las ingentes cantidades de información requeridas para el entrenamiento, lo cual supone un proceso lento y costoso.

Deep Learning:

El Deep Learning lleva a cabo el proceso de Machine Learning usando una red neuronal artificial que se compone de un número de niveles jerárquicos. En el nivel inicial de la jerarquía la red aprende algo simple y luego envía esta información al siguiente nivel. El siguiente nivel toma esta información sencilla, la combina, compone una información algo un poco más compleja, y se lo pasa al tercer nivel, y así sucesivamente.

Continuando con el ejemplo del gato, el nivel inicial de una red de Deep Learning podría utilizar las diferencias entre las zonas claras y oscuras de una imagen para saber dónde están los bordes de la imagen. El nivel inicial pasa esta información al segundo nivel, que combina los bordes construyendo formas simples, como una línea diagonal o un ángulo recto. El tercer nivel combina las formas simples y obtiene objetos más complejos como óvalos o rectángulos. El siguiente nivel podría

combinar los óvalos y rectángulos, formando barbas, patas o colas rudimentarias. El proceso continúa hasta que se alcanza el nivel superior en la jerarquía, en el cual la red aprende a identificar gatos.

Aplicaciones de Deep Learning:

El Deep Learning ha llamado mucho la atención por su potencial utilidad en distintos tipos de aplicaciones en el "mundo real" (pueden aplicarse con éxito a grandes volúmenes de datos para el descubrimiento y aplicación de conocimiento, así como a la realización de predicciones a partir de él), principalmente debido a que obtiene tasas de éxito elevadas con entrenamiento "no supervisado". En el caso del ejemplo, las redes de Deep Learning aprenderían a identificar gatos aunque las imágenes no tuvieran la etiqueta "gato".

A continuación se muestran algunos de los principales problemas "reales" en los que distintas compañías están aplicando Deep Learning en la actualidad:

- Utilización de imágenes en lugar de palabras clave para buscar productos de una empresa, o artículos similares.
- Identificar marcas y logotipos de empresas en fotos publicadas en redes sociales.
- Monitorización en tiempo real de reacciones en canales online durante el lanzamiento de productos.
- Orientación de anuncios y predicción de las preferencias de los clientes.
- Identificación y seguimiento de los niveles de confianza de los clientes, sus opiniones y actitud en diferentes canales online y servicios de soporte automatizado al cliente.
- Identificación de clientes potenciales.
- Detección de fraudes, recomendaciones a clientes, gestión de relaciones con los clientes, etc.
- Mejor comprensión de enfermedades, mutaciones de enfermedades y terapias genéticas.
- Análisis de imágenes médicas, como radiografías y resonancias magnéticas, aumentando la precisión diagnóstica, en un menor tiempo y con un menor coste que los métodos tradicionales.
- Exploración de la posibilidad de reutilización de fármacos ya conocidos y probados para su uso contra nuevas enfermedades.
- Identificación en textos de sentimientos positivos y negativos, temas y palabras clave.

- Detección, predicción y prevención de amenazas sofisticadas en tiempo real en el campo de la ciberseguridad.
- Localización de caras e identificación de emociones faciales.
- Reconocimiento de voz.
- Clasificación de vídeos.

[7]

<http://www.tecnicaindustrial.es/TIFrontal/a-625-conceptos-aplicaciones-logica-borrosa.aspx>

[8]

<http://www.uco.es/~i42crjj/aplicada/tema4.htm>

[9]

<http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/temageneticos.pdf>

[10]

<https://www.indracompany.com/es/blogneo/deep-learning-sirve>

[11]

<https://www.smartpanel.com/que-es-deep-learning/>

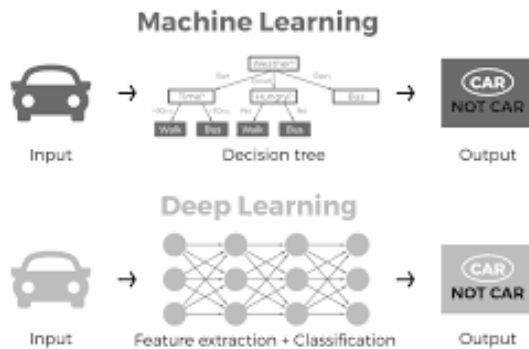


Fig.8 Diferencia entre Machine Learning y Deep Learning

REFERENCIAS

Referencias en la Web:

[1]

<https://computerhoy.com/reportajes/tecnologia/inteligencia-artificial-469917>

[2]

https://www.ecured.cu/Redes_neuronales_artificiales

[3]

<https://www.atriainnovation.com/que-son-las-redes-neuronales-y-sus-funciones/>

[4]

https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5_anio/orientadora1/monograias/matich-redesneuronales.pdf

[5]

<https://sites.google.com/site/logicadifusaingindustria/alpaita/logica-difusa/ventajas-y-desventajas-de-la-logica-difusa>

[6]

<https://www.tdx.cat/bitstream/handle/10803/6887/0/4Rpp04de11.pdf>