Resumen Inteligencia Artificial 2

Por: Pablo Federico Martín Luna

# Tema 1: Introducción al Procesamiento Subsimbólico

## Inteligencia Artificial Subsimbólico

Técnicas de IA que simulan los elementos de más bajo nivel que interviene en los procesos inteligentes para que de su combinación emerja de forma espontánea el comportamiento inteligente.

### Soft Computing

Rama de la IA que se dedica al desarrollo de algoritmos y técnicas que pueden resolver problemas que son típicamente difíciles o no factibles de resolver utilizando los métodos tradicionales. (No tiene por qué ser la mejor solución)

### Técnicas subsimbólicas basadas en sistemas biológicos

* Estructura del cerebro: Redes Neuronales
* Evolución biológica: Computación evolutiva

### Áreas

* Búsqueda
* Aprendizaje automático (Machine Learning/Data Mining)
* Simulación sensorial

## Machine Learning (Aprendizaje Automático

Rama de la IA que desarrolla algoritmos de análisis de datos que permitan al sistema aprender y tomar decisiones autónomas en base a la predicción de las situaciones que se puedan producir.

### Aprender

Generalizar comportamientos mediante detección de patrones en la información suministrada en forma de ejemplos y experiencia.

### Optimizar

Calcular el máximo y mínimo de una función.

Función de error, Loss function, Función objetivo.

En machine Learning ha aprendido porque nuestra función de error es un mínimo.

### Factores que influyen en el aprendizaje

* Qué **conocimiento previo** tiene el sistema. (un buen set de datos)
* Qué **representación** se usa para los datos que usan el sistema.
* Qué **feedback** está disponible para que el sistema aprenda.
  + Aprendizaje supervisado:
    - Cualitativa – clasificación.
    - Cuantitativa – regresión.
  + Aprendizaje no supervisado:
    - Clustering
  + Aprendizaje por refuerzo.
  + Aprendizaje semi-supervisado

Interfaz de usuario gráfica

Descripción generada automáticamente con confianza media

## ¿Qué son las Redes Neuronales?

Sistema compuesto por un conjunto de elementos de procesamiento simples (neuronas), conectadas entre sí (red), cuyo comportamiento está determinado por la topología y los pesos de sus conexiones.

## Características

### Propiedades emergentes

El comportamiento global de la red es distinto del comportamiento individual de las neuronas que la componen. (aprender de la experiencia, autoorganizarse, generalizar y abstraer)

### Aprendizaje adaptativo

El sistema modifica su comportamiento para ajustarlo a los datos que se le proporcionan.

### Autoorganización

Emplean su capacidad de aprendizaje adaptativo para autoorganizar la información que recibe durante el aprendizaje.

### Generalización

Facultad de responder apropiadamente cuando se les presentan datos que no habían sido presentados anteriormente.

### Procesamiento distribuido

El conocimiento no está en ninguna neurona concreta.

## Conceptos básicos

### Neurona

Unidad de procesamiento elemental que genera una salida como resultado de la integración de un conjunto de entradas.

Formada por: entradas, salidas, procesamiento y pesos.

* Procesamiento: operaciones elementales, suma ponderada, funciones de activación.
  + Funciones de activación:
    - No lineal:

Imagen que contiene Escala de tiempo

Descripción generada automáticamente

* + - Lineal:

Gráfico, Diagrama, Gráfico de líneas

Descripción generada automáticamente

### Capa

Elemento estructural que agrupa neuronas con comportamiento similar.

### Red

Conjunto de reglas que define la forma en que se conectan las capas y/o PE.

Multicapa: Capa de Entrada (IL), Capa/s intermedia/s (HL) y Capa de Salida (OL).

### Procesamiento

A partir de la capa de entrada cada neurona

1. Procesa la información de sus conexiones de entrada.
2. Evalúa su función de activación.
3. Transmite su salida a las neuronas con las que está conectada.

### Aprendizaje

Proceso (algoritmo) por el que la red modifica sus pesos en función de las salidas presentadas asociando patrones de entrada a patrones de salida.

## Tipos de arquitecturas

Diagrama

Descripción generada automáticamente

# Tema 2: Aprendizaje NO Supervisado: Aprendizaje Competitivo

Algoritmos de aprendizaje no supervisado están generalmente relacionados con la detección de características: Clustering y Clasificar.

## Clustering

Dividir el espacio N-dimensional de características en K regiones excluyentes.

Las características N-dimensionales deben ser multinormales con igual covarianza.

## Clasificar

Asignar una de las K regiones excluyentes obtenidas en el Clustering a los patrones.

Asignar cada patrón a uno de los grupos/clases definidos Previamente.

## Tipos de clasificadores

### Discriminación

* Discriminantes lineales o no lineales.
  + Clasificación supervisada.
  + Parten de los valores de una serie limitada de parámetros.
  + Comparación de características del individuo y de los grupos.
* Regresión logística.

### Clustering

* Métodos de Clustering o estimación de densidades.
  + Clasificación no supervisada.
  + No se dispone de una muestra previamente clasificada. A priori no se conocen los grupos y lo que precisamente se desea es establecerlos a partir de los datos que poseemos.
  + La agrupación de individuos se realiza minimizando ciertas funciones de distancia mediante técnicas estadísticas:
    - Distancia Euclídea:
    - Distancia de Manhattan:
    - Distancia de Mahalanobis.

## Autoorganización

Proceso en el cual, por medio de interacciones locales, se obtiene un ordenamiento global mediante un proceso de aprendizaje no supervisado.

## Tipos de aprendizaje no supervisado

### Aprendizaje Hebbiano o Asociativo

1. Si dos neuronas en cada lado de una conexión se activan sincrónicamente, entonces el peso de esa conexión se incrementa.
2. Si dos neuronas en cada lado de una conexión se activan asincrónicamente, entonces el peso de esa conexión se decrementa.

### Aprendizaje por Principal Component Analysis (PCA)

(Esto lo estudiamos del Resumen Data Mining)

### Aprendizaje Competitivo (SOM y ART)

El aprendizaje consiste en reforzar las conexiones de la unidad ganadora y debilitar las otras.

## Mapas Autoorganizativos

### Arquitectura

* Capas: Entrada + Kohonen.
* Elementos de un SOM:
  + Matriz multidimensional.
  + Función discriminante.
  + Mecanismo de competición.
  + Mecanismo de cooperación.
  + Mecanismo de adaptación.

### Procesamiento

1. Propagación de la entrada.
2. Cálculo de la función discriminante.

1. Competición.

### Aprendizaje

Ajuste de la matriz de pesos de las neuronas de la capa de Kohonen a medida que se presentan patrones (vectores) de entrada en un proceso iterativo (entrenamiento).

Forma

Descripción generada automáticamente

Proceso:

1. Se inicializa la matriz de pesos con valores aleatorios.
2. Se presentan uno a uno los patrones del conjunto de entrenamiento, también normalizados.
3. Se determina la neurona ganadora (BMU).
4. Se actualizan los pesos de la BMU.

Vecindario

El aprendizaje no se limita a la BMU, sino que se extiende a las neuronas adyacentes.

Las funciones de aprendizaje y vecindario están ligadas.

### Mapas topológicos

La capa de Kohonen preserva el orden y compacta la representación de una nube de datos en un espacio N-dimensional proyectándola sobre un mapa bidimensional.

### Parámetros

* Periodo. (p)
* Coeficiente de aprendizaje. ()
* Vecindario. ()
* Amortiguación. ()
* Número de neuronas del lado de la capa de Kohonen. (k)

Iteración: Activación de una neurona + Aprendizaje.

Número de presentaciones. (t)

Variación del coeficiente de aprendizaje.

Variación del vecindario.

Amortiguación.

Ecuación de aprendizaje.

### Modificaciones al proceso básico

* Refuerzo
  + Después del primer periodo de aprendizaje, se realizan sucesivos períodos de refuerzo.
  + En cada periodo de refuerzo se cumple que:

* Mecanismo de consciencia
  + Función de penalización respecto a la frecuencia de activación.
  + Esto se realiza modificando la distancia neurona – patrón a partir de la frecuencia de activación de cada neurona.

## Métricas para medir la adecuación de cada mapa

* Error de Cuantificación.

Distancia promedio entre cada vector de datos y su BMU.

* Distancia media de cada clase.
* Distancia media de la red.
* Error Topológico.

Proporción de todos los vectores de datos para los que la primera y la segunda BMU no son unidades adyacentes.

* Primera y segunda BMU adyacentes – 0.
* Primera y segunda BMU no adyacentes – 1.