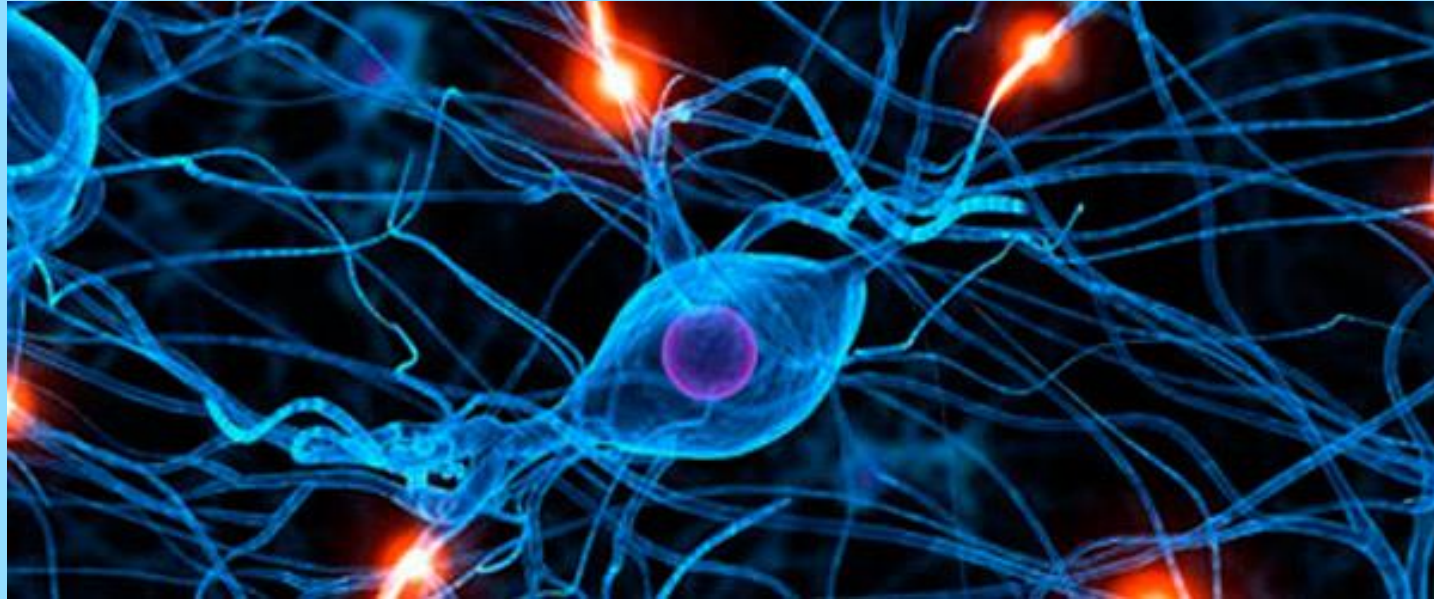


REDES NEURONALES APLICADAS A LA MINERÍA DE DATOS



MINERÍA DE DATOS – REDES NEURONALES

- ❑ INTRODUCCIÓN.
- ❑ DESCRIPCIÓN DE UNA RED NEURONAL (RN).
- ❑ APRENDIZAJE EN LAS RN.
- ❑ FUNCIONAMIENTO DE UNA RN.
- ❑ EL ALGORITMO DE APRENDIZAJE DE RETRO-PROPAGACIÓN.
- ❑ ANÁLISIS DISCRIMINANTE A TRAVÉS DEL PERCEPTRÓN.
- ❑ ANÁLISIS DE SERIES TEMPORALES MEDIANTE RN.
- ❑ ANÁLISIS DE COMPONENTES PRINCIPALES CON RN.
- ❑ AGRUPAMIENTO (CLUSTERING) MEDIANTE RN.

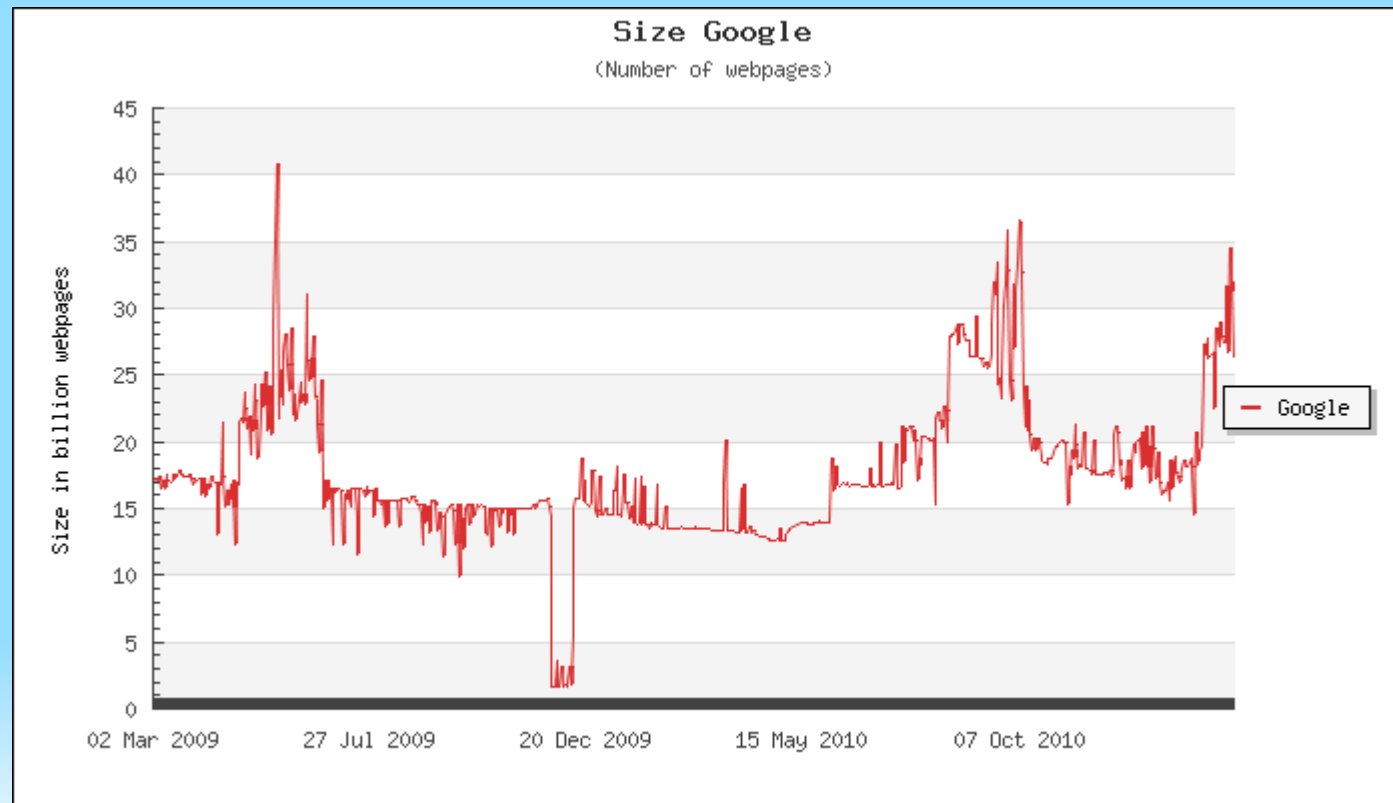
INTRODUCCIÓN

INTRODUCCIÓN

- ❑ **¿CUÁL ES EL PROBLEMA?**
- ❑ **LA CRECIENTE BRECHA ENTRE LA GENERACIÓN DE INFORMACIÓN Y NUESTRA CAPACIDAD PARA ENTENDERLA.**
- ❑ **NO SE ESTÁ APROVECHANDO LA INFORMACIÓN POTENCIALMENTE ÚTIL QUE SE ENCUENTRA “ESCONDIDA” EN LOS DATOS.**

INTRODUCCIÓN

- ❑ **EL TAMAÑO DE LA WWW: TAMAÑO ESTIMADO DE LOS ÍNDICES DE GOOGLE**
- ❑ **LA WEB INDEXADA CONTIENE MÁS DE 35 BILLONES DE PÁGINAS.**

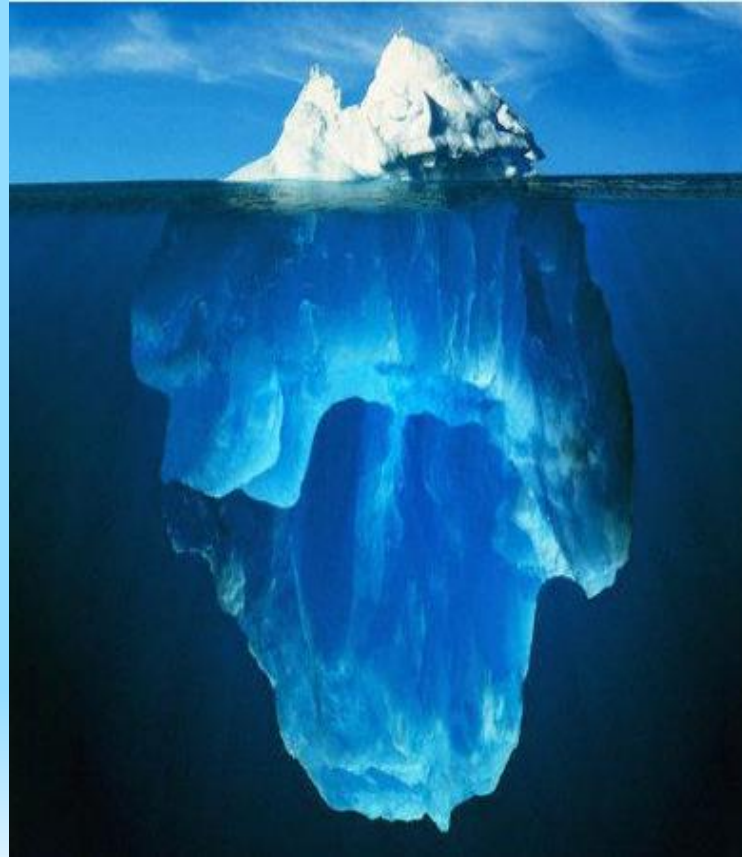


INTRODUCCIÓN

- ❑ **NUEVAS DISCIPLINAS**
- ❑ **MINERÍA DE DATOS: DM.**
- ❑ **DESCUBRIMIENTO DE *CONOCIMIENTO* EN BASES DE DATOS: KDD.**
- ❑ **MINERÍA DE DATOS:**
 - ❖ **TÉCNICAS QUE PERMITEN EXTRAER INFORMACIÓN ÚTIL QUE ESTABA IMPLÍCITA, Y POR TANTO DESCONOCIDA, EN LOS DATOS.**
 - ❖ **BASADA EN:**
 - **ESTADÍSTICA.**
 - **APRENDIZAJE DE MÁQUINA.**
 - **TEORÍA DE LA INFORMACIÓN.**
 - **CIENCIAS DE LA COMPUTACIÓN.**

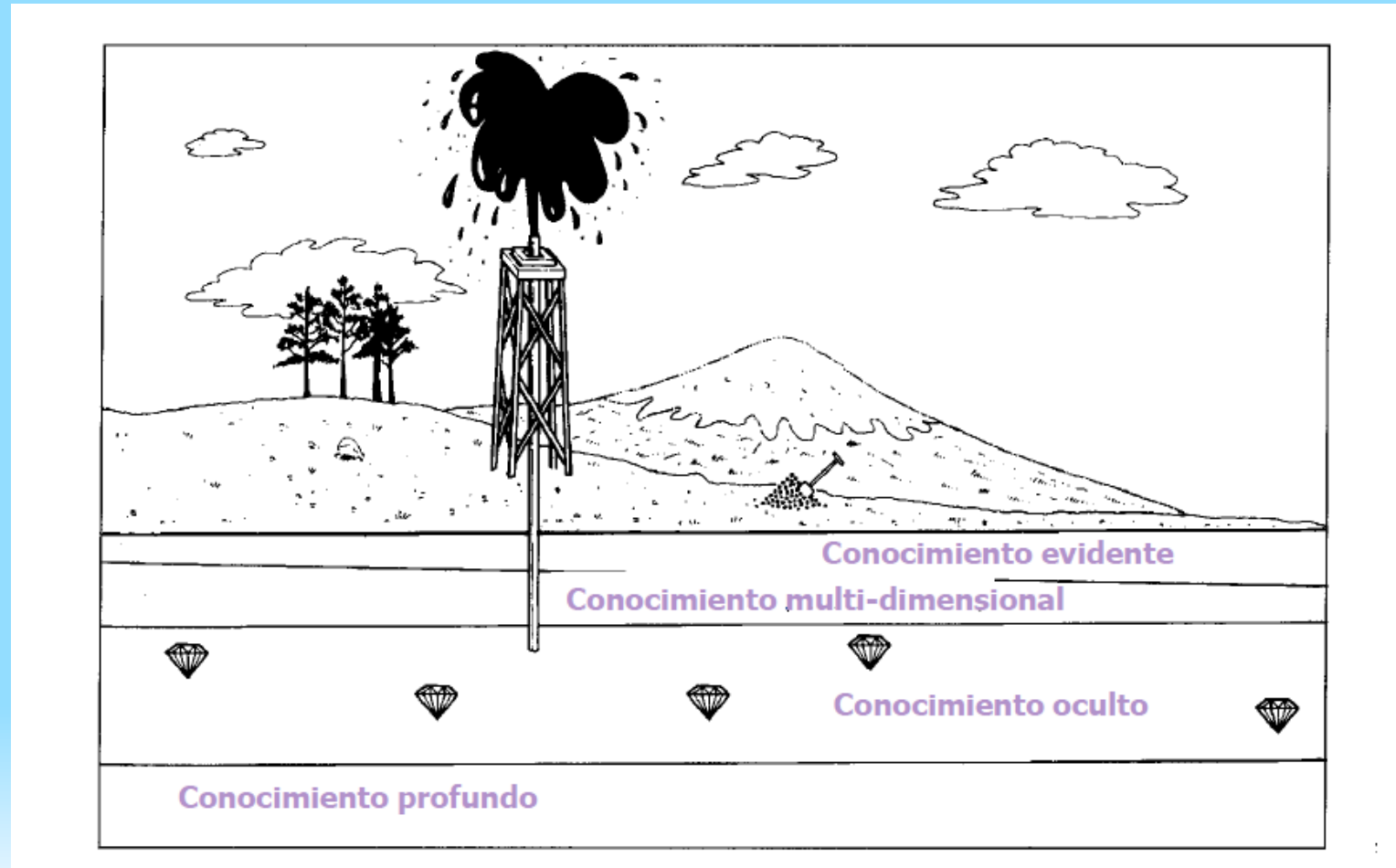
INTRODUCCIÓN

- ❑ ¿CUÁL ES EL GRAN DESAFÍO?



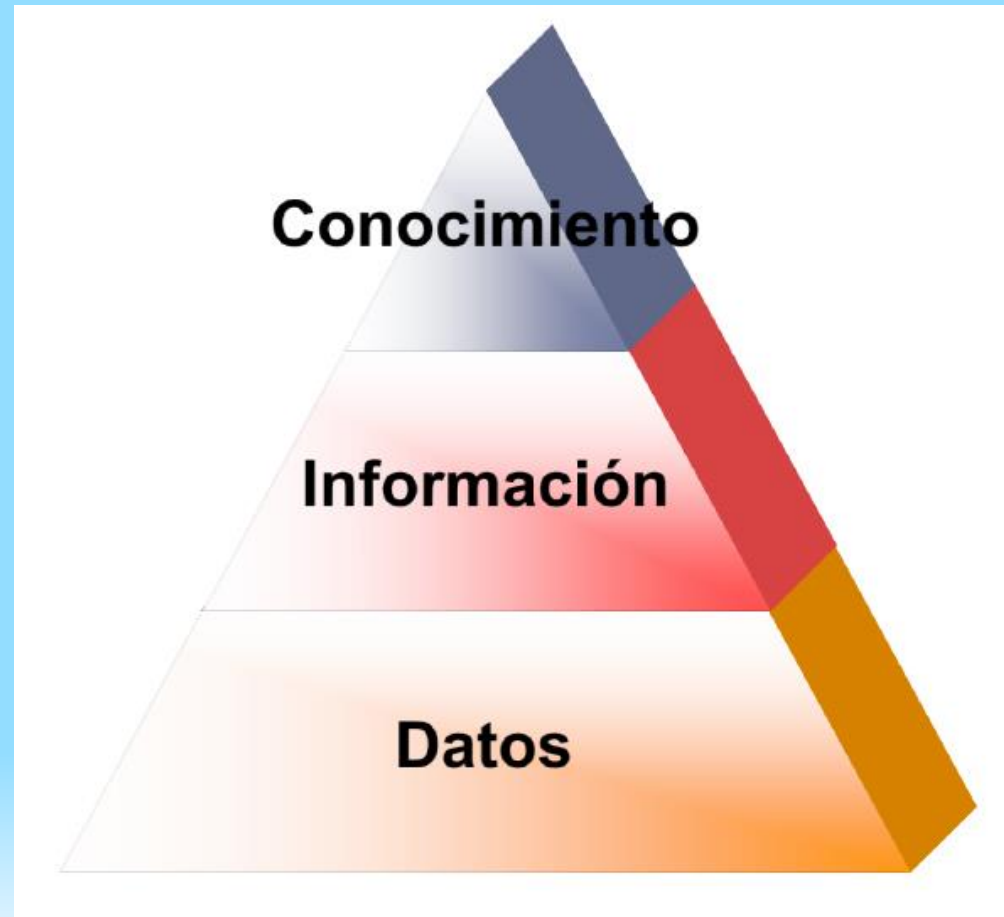
INTRODUCCIÓN

❑ TIPOS DE CONOCIMIENTO



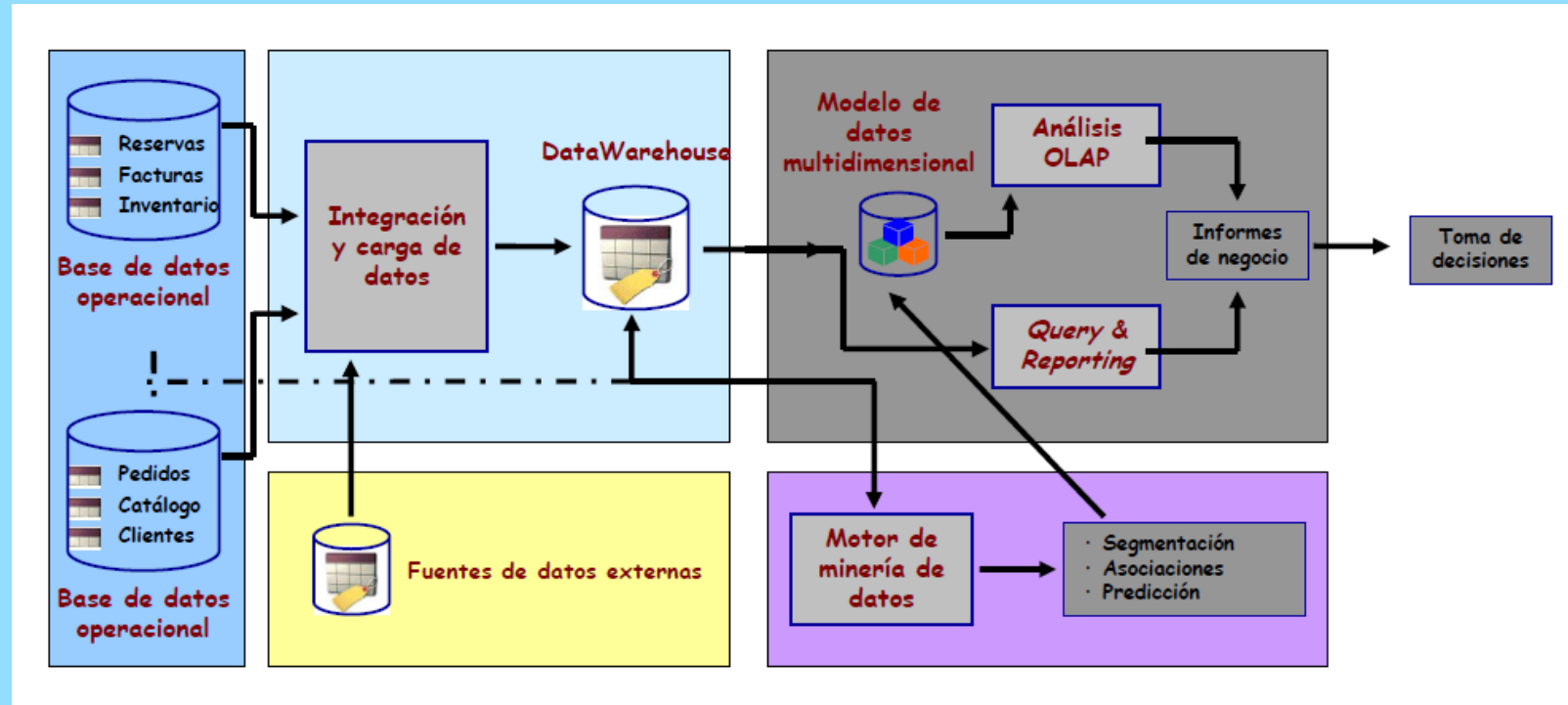
INTRODUCCIÓN

❑ DATOS, INFORMACIÓN Y CONOCIMIENTO



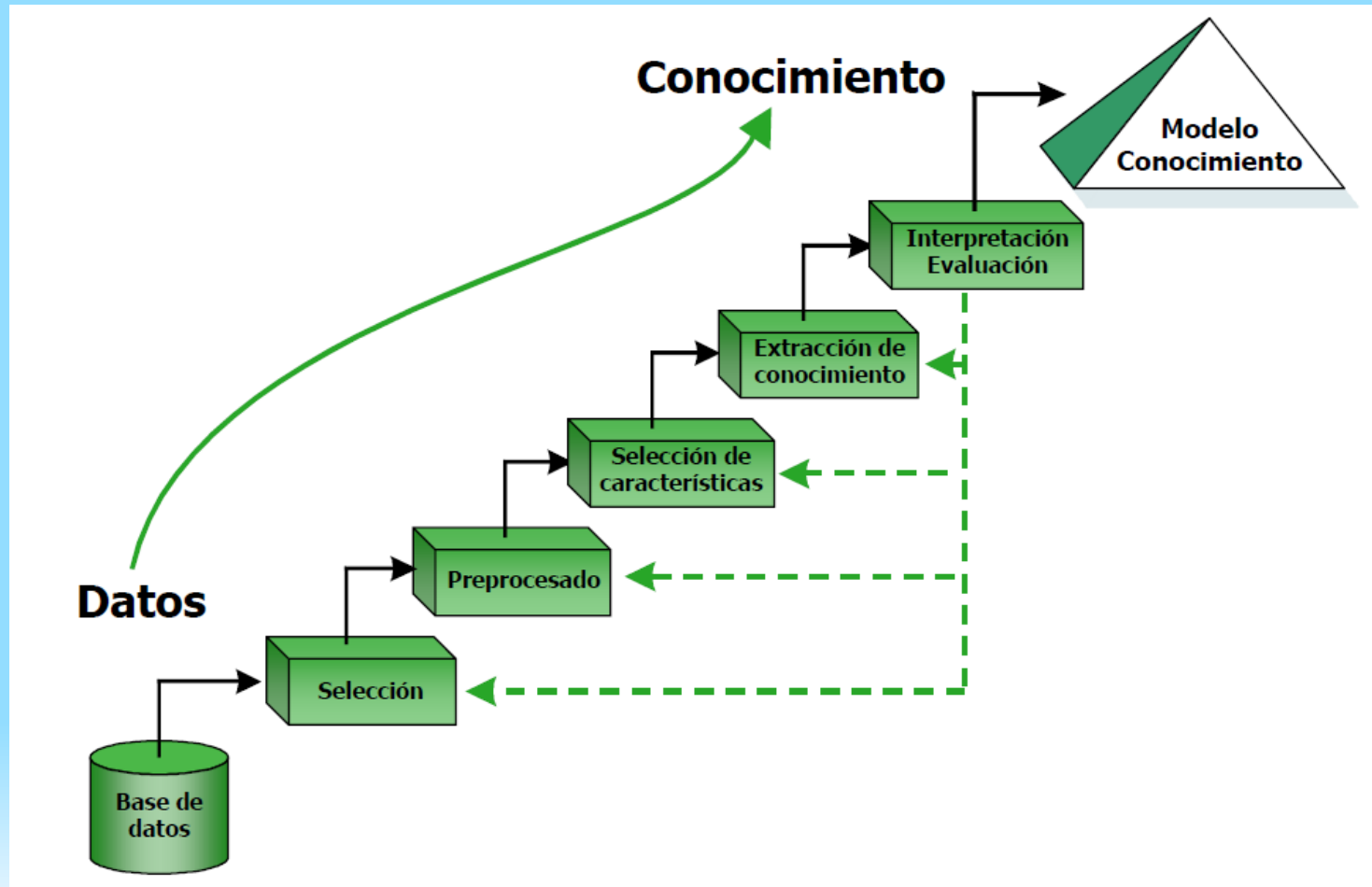
INTRODUCCIÓN

❑ ARQUITECTURA TECNOLÓGICA



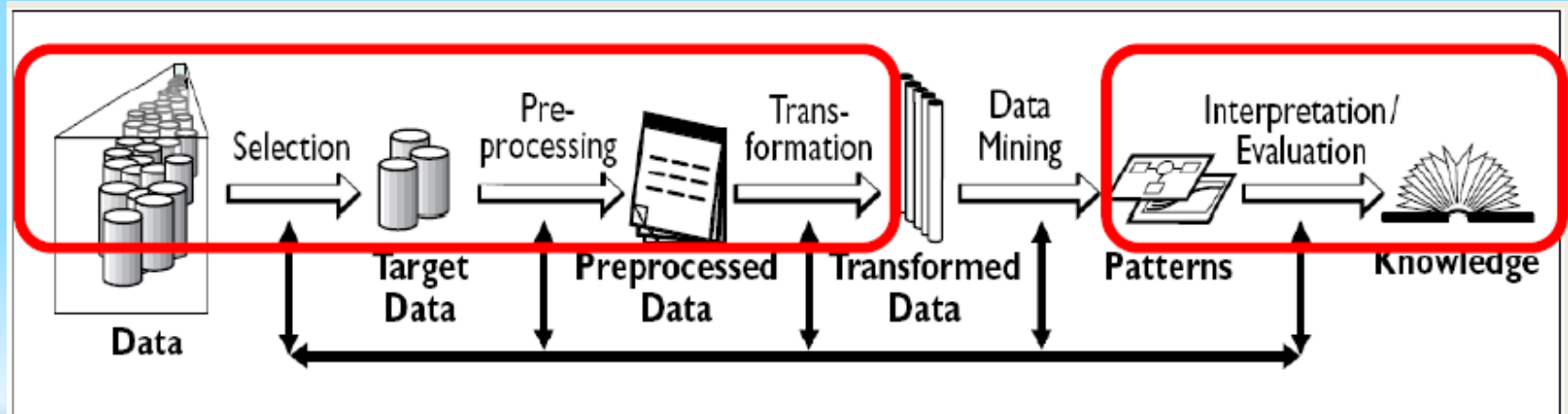
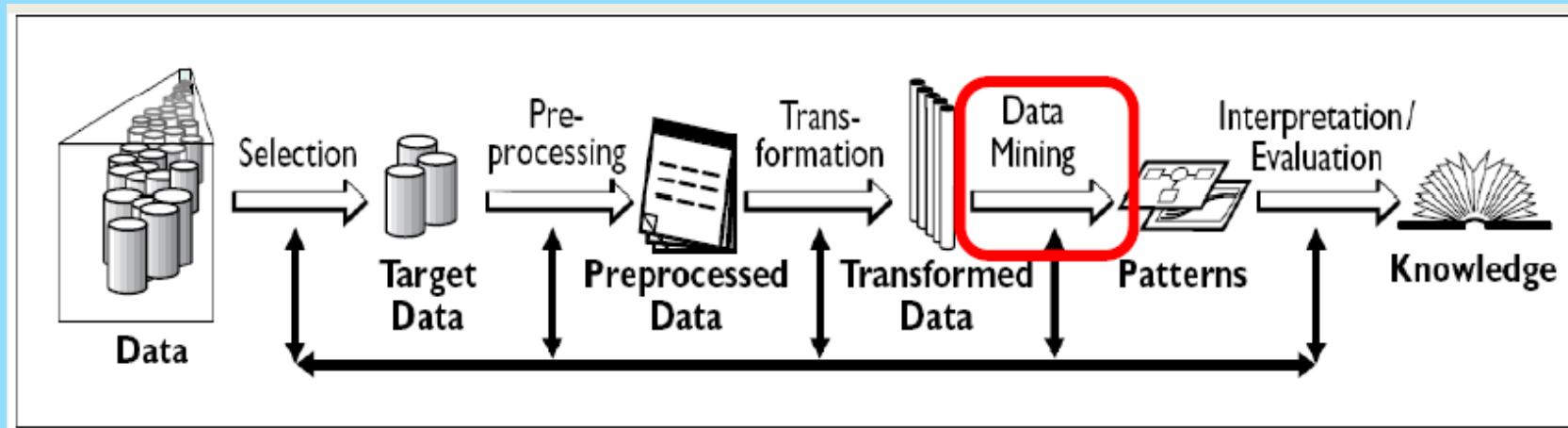
INTRODUCCIÓN

❑ PROCESO DE MINERÍA DE DATOS



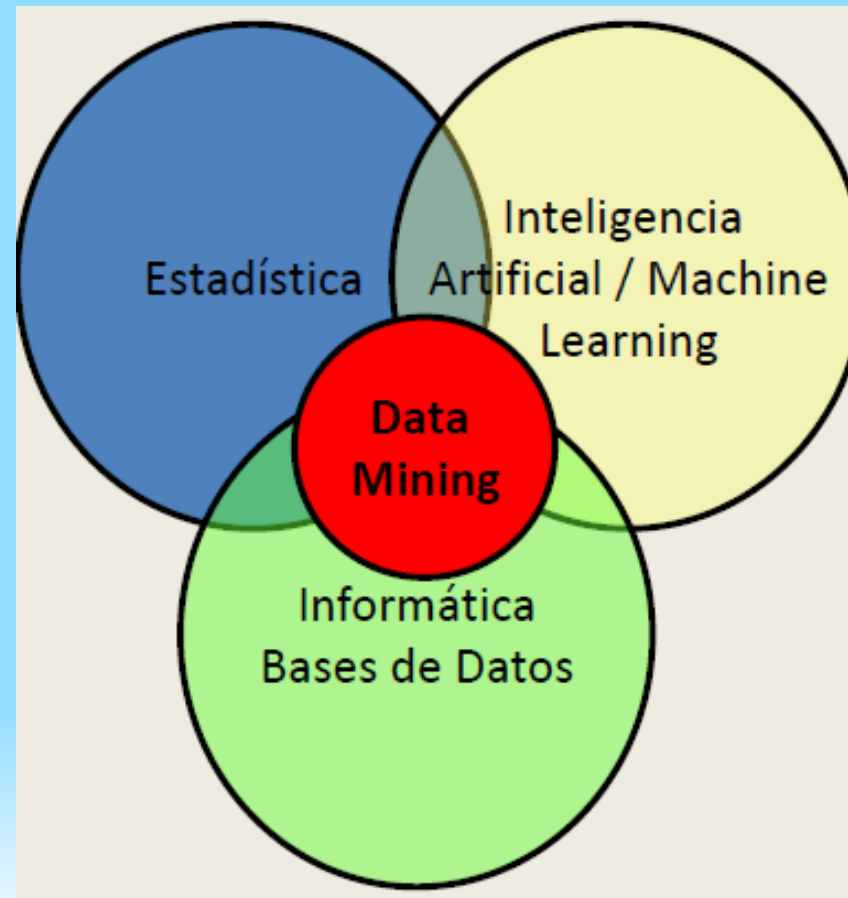
INTRODUCCIÓN

- ❑ **PROCESO DE MINERÍA DE DATOS COMO PARTE DEL PROCESO DE DESCUBRIMIENTO DE CONOCIMIENTO (KDD)**



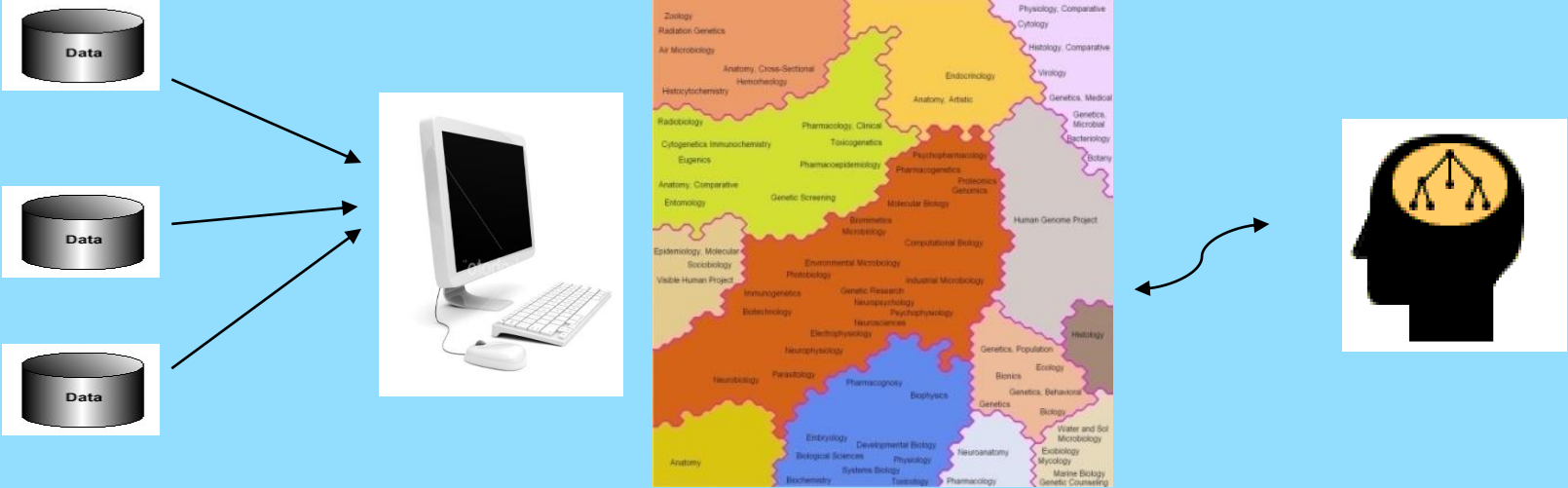
INTRODUCCIÓN

- ❑ **PROCESO DE MINERÍA DE DATOS:**
 - ❖ ¿QUÉ HERRAMIENTAS UTILIZA?



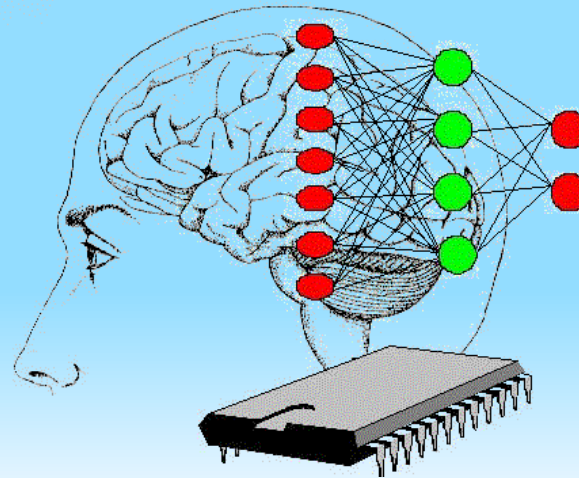
INTRODUCCIÓN

❑ EXTRACCIÓN AUTOMÁTICA DE PATRONES



INTRODUCCIÓN

- ❑ **MINERÍA DE DATOS**
- ❑ **EXTRAER AUTOMÁTICAMENTE “MODELOS” A PARTIR DE DATOS:**
 - ❖ **ÁRBOLES DE DECISIÓN.**
 - ❖ **REGLAS DE ASOCIACIÓN.**
 - ❖ **AJUSTE DE FUNCIONES.**
 - ❖ **REDES NEURONALES ARTIFICIALES.**
 - ❖ **REDES BAYESIANAS.**
 - ❖ **ETC.**



DESCRIPCIÓN DE UNA RED NEURONAL (RN)

DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **DEFINICIÓN**
- ❑ **ES UN CONJUNTO DE ELEMENTOS DE PROCESAMIENTO DE LA INFORMACIÓN ALTAMENTE INTERCONECTADOS, QUE SON CAPACES DE APRENDER CON LA INFORMACIÓN QUE SE LES ALIMENTA.**
- ❑ **PUEDE APLICARSE A GRAN NÚMERO DE PROBLEMAS QUE PUEDEN IR DESDE PROBLEMAS COMPLEJOS REALES A MODELOS TEÓRICOS SOFISTICADOS, COMO POR EJ.:**
 - ❖ RECONOCIMIENTO DE IMÁGENES.
 - ❖ RECONOCIMIENTO DE VOZ.
 - ❖ ANÁLISIS Y FILTRADO DE SEÑALES.
 - ❖ CLASIFICACIÓN.
 - ❖ DISCRIMINACIÓN.
 - ❖ ANÁLISIS FINANCIERO.
 - ❖ PREDICCIÓN DINÁMICA, ETC.

DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **DEFINICIÓN**
- ❑ **LAS RN:**
 - ❖ **TRATAN DE EMULAR EL SISTEMA NERVIOSO.**
 - ❖ **SON CAPACES DE REPRODUCIR ALGUNAS DE LAS TAREAS QUE DESARROLLA EL CEREBRO HUMANO.**
 - ❖ **REFLEJAN ALGUNAS CARACTERÍSTICAS FUNDAMENTALES DE COMPORTAMIENTO DEL CEREBRO.**
- ❑ **INTENTAN MODELIZAR:**
 - ❖ **UNA DE LAS ESTRUCTURAS FISIOLÓGICAS DE SOPORTE DEL CEREBRO, LA NEURONA.**
 - ❖ **LOS GRUPOS ESTRUCTURADOS E INTERCONECTADOS DE VARIAS DE ELLAS, CONOCIDOS COMO REDES DE NEURONAS.**
- ❑ **CONSTRUYEN SISTEMAS QUE PRESENTAN UN CIERTO GRADO DE INTELIGENCIA.**

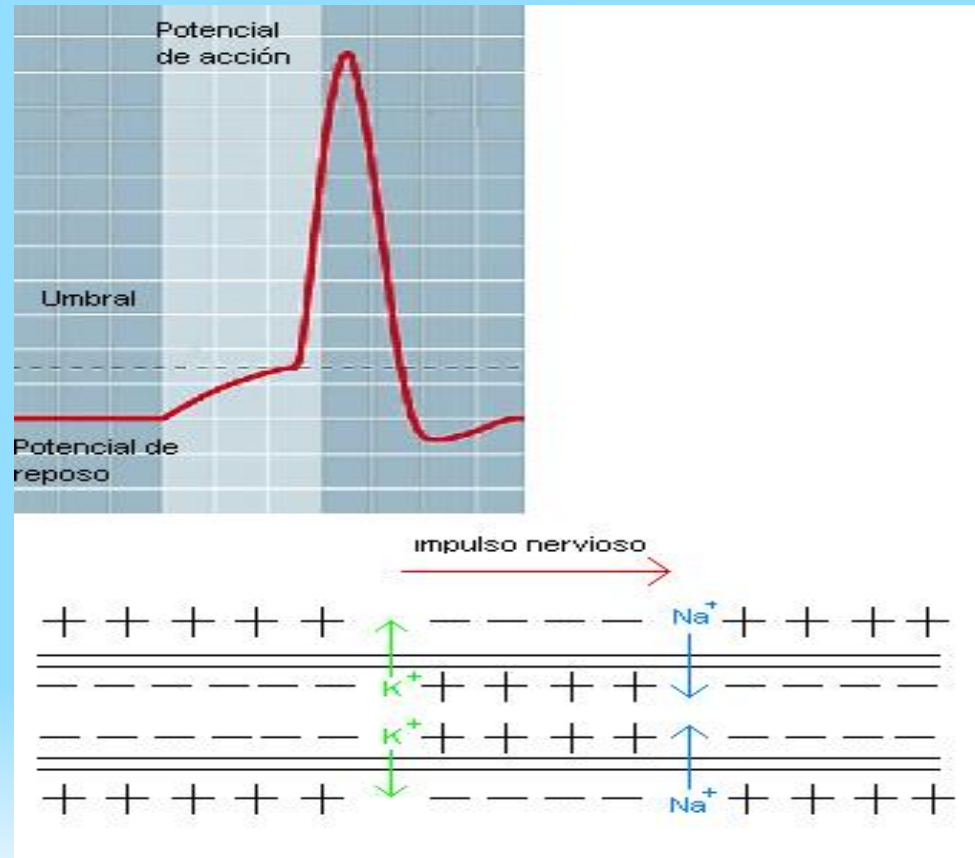
DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **DEFINICIÓN**
- ❑ **TEORÍA DE LA NEURONA**
- ❑ **ESTRUCTURA BÁSICA Y FUNCIONAL DEL SISTEMA NERVIOSO:**
 - ❖ **SANTIAGO RAMÓN Y CAJAL.**



DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ DEFINICIÓN
- ❑ TEORÍA DE LA NEURONA



DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **DEFINICIÓN**
- ❑ **LOS SISTEMAS NEURONALES ARTIFICIALES TIENEN LIMITACIONES Y SÓLO POSEEN UN PARECIDO SUPERFICIAL CON SUS CONTRAPARTIDAS BIOLÓGICAS.**
- ❑ **LAS RN, EN RELACIÓN CON EL PROCESAMIENTO DE INFORMACIÓN, HEREDAN TRES CARACTERÍSTICAS BÁSICAS DE LAS REDES DE NEURONAS BIOLÓGICAS:**
 - ❖ **PARALELISMO MASIVO.**
 - ❖ **RESPUESTA NO LINEAL DE LAS NEURONAS FRENTE A LAS ENTRADAS RECIBIDAS.**
 - ❖ **PROCESAMIENTO DE INFORMACIÓN A TRAVÉS DE MÚLTIPLES CAPAS DE NEURONAS.**

DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **DEFINICIÓN**
- ❑ UNA DE SUS PRINCIPALES **PROPIEDADES** ES SU CAPACIDAD DE **APRENDER Y GENERALIZAR A PARTIR DE EJEMPLOS REALES**:
 - ❖ LA RED **APRENDE A RECONOCER LA RELACIÓN QUE EXISTE ENTRE**:
 - EL CONJUNTO DE **ENTRADAS** PROPORCIONADAS COMO EJEMPLOS, Y.
 - SUS CORRESPONDIENTES **SALIDAS**.
 - ❖ **FINALIZADO EL APRENDIZAJE**:
 - CUANDO A LA RED SE LE PRESENTA UNA **NUEVA ENTRADA** (AUNQUE ESTÉ INCOMPLETA O POSEA ALGÚN ERROR).
 - ES CAPAZ DE GENERALIZARLA OFRECIENDO UNA **SALIDA**, EN BASE A LA **RELACIÓN FUNCIONAL** ESTABLECIDA EN EL APRENDIZAJE.

DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **DEFINICIÓN**
- ❑ **SE PUEDE DEFINIR UNA RN ARTIFICIAL COMO UN SISTEMA INTELIGENTE CAPAZ:**
 - ❖ **APRENDER.**
 - ❖ **GENERALIZAR.**
- ❑ **ESTÁ FORMADA POR UNIDADES DE PROCESAMIENTO QUE RECIBEN EL NOMBRE DE NEURONAS O NODOS:**
 - ❖ **ESTÁN ORGANIZADOS EN GRUPOS QUE SE LLAMAN “CAPAS”.**

DESCRIPCIÓN DE UNA RED NEURONAL (RN)

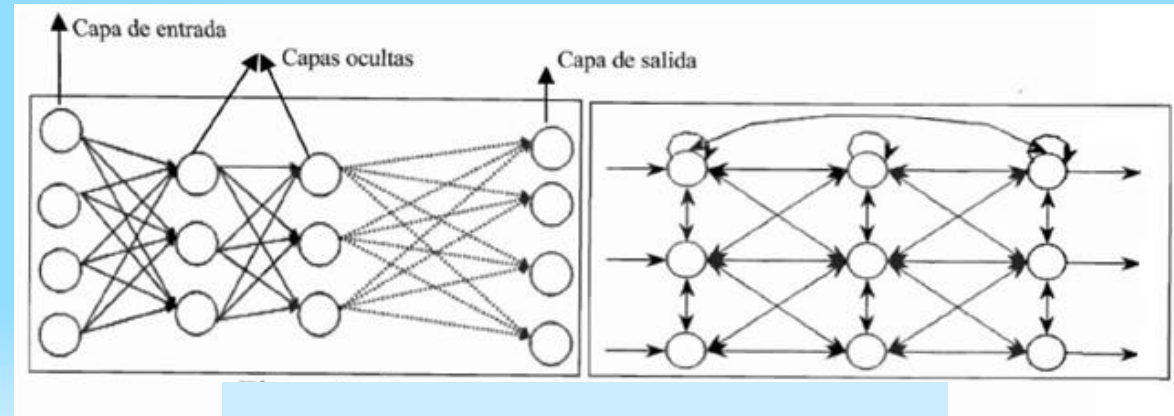
- ❑ **DEFINICIÓN**
- ❑ **GENERALMENTE EXISTEN TRES TIPOS DE CAPAS:**
 - ❖ **UNA CAPA DE ENTRADA.**
 - ❖ **UNA O VARIAS CAPAS OCULTAS.**
 - ❖ **UNA CAPA DE SALIDA.**
- ❑ **LAS CONEXIONES SE ESTABLECEN ENTRE LOS NODOS DE CADA CAPA ADYACENTES.**
- ❑ **LA CAPA DE ENTRADA ESTÁ FORMADA POR NODOS DE ENTRADA QUE RECIBEN LA INFORMACIÓN DIRECTAMENTE DEL EXTERIOR.**
- ❑ **LA CAPA DE SALIDA REPRESENTA LA RESPUESTA DE LA RED A UNA ENTRADA DADA SIENDO ESTA INFORMACIÓN TRANSFERIDA AL EXTERIOR.**

DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **DEFINICIÓN**
- ❑ **LAS CAPAS OCULTAS O INTERMEDIAS:**
 - ❖ **SE ENCARGAN DE PROCESAR LA INFORMACIÓN.**
 - ❖ **SE INTERPONEN ENTRE LAS CAPAS DE ENTRADA Y SALIDA.**
 - ❖ **SON LAS ÚNICAS QUE NO TIENEN CONEXIÓN CON EL EXTERIOR.**
- ❑ **LA ESTRUCTURA MÁS HABITUAL ES LA DENOMINADA RED ALIMENTADA HACIA DELANTE O FEEDFORWARD:**
 - ❖ **LAS CONEXIONES ENTRE NEURONAS SE ESTABLECEN EN UN ÚNICO SENTIDO, POR EL SIGUIENTE ORDEN:**
 - **CAPA DE ENTRADA.**
 - **CAPA(S) OCULTA(S).**
 - **CAPA DE SALIDA.**

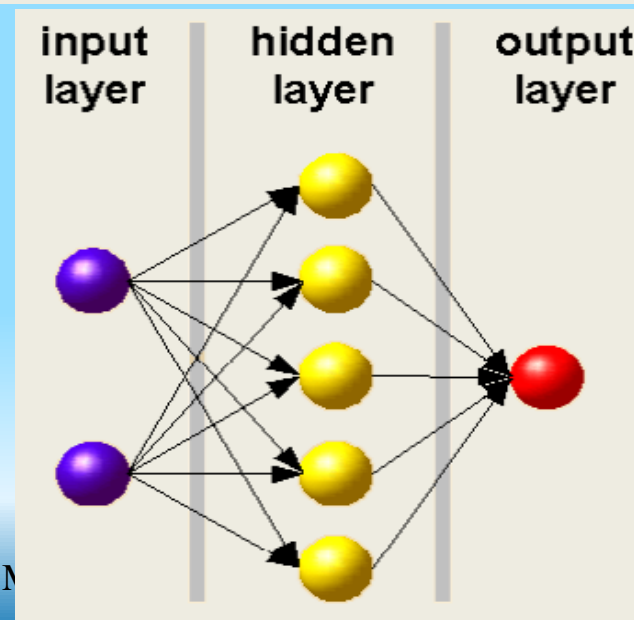
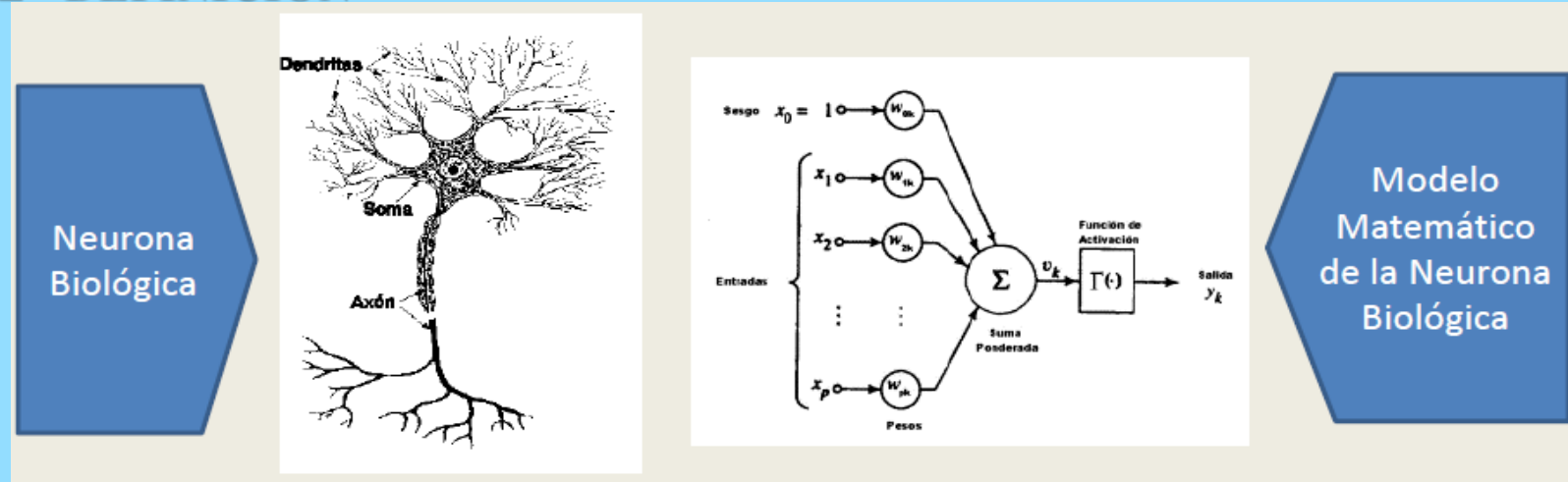
DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **DEFINICIÓN**
- ❑ **EXISTEN TAMBIÉN REDES RETROALIMENTADAS O FEEDBACK:**
 - ❖ **PUEDEN TENER CONEXIONES HACIA ATRÁS.**
 - ❖ **DE NODOS DE UNA CAPA A ELEMENTOS DE PROCESO DE CAPAS ANTERIORES.**
- ❑ **TAMBIÉN HAY REDES RECURRENTE:**
 - ❖ **PUEDEN POSEER CONEXIONES, TANTO ENTRE NEURONAS DE UNA MISMA CAPA, COMO DE UN NODO A SÍ MISMO.**



DESCRIPCIÓN DE UNA RED NEURONAL (RN)

▣ DEFINICIÓN

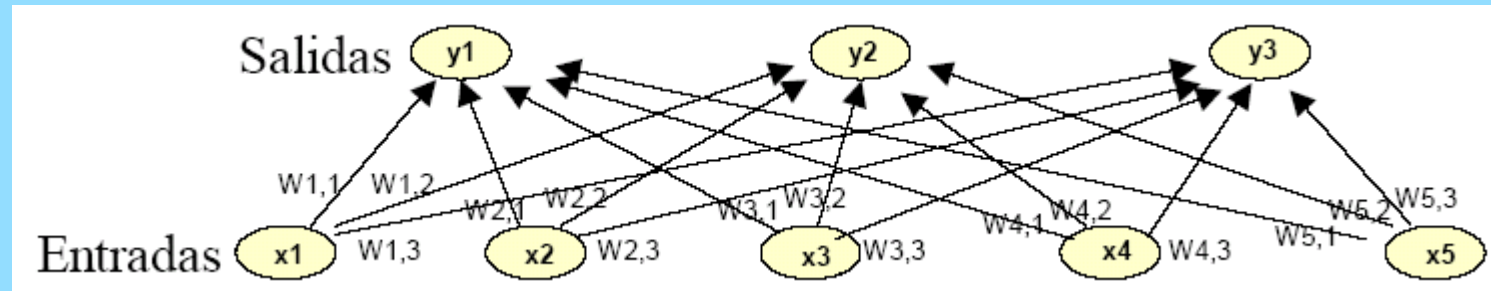


DESCRIPCIÓN DE UNA RED NEURONAL (RN)

❑ PERCEPTRON LEARNING:

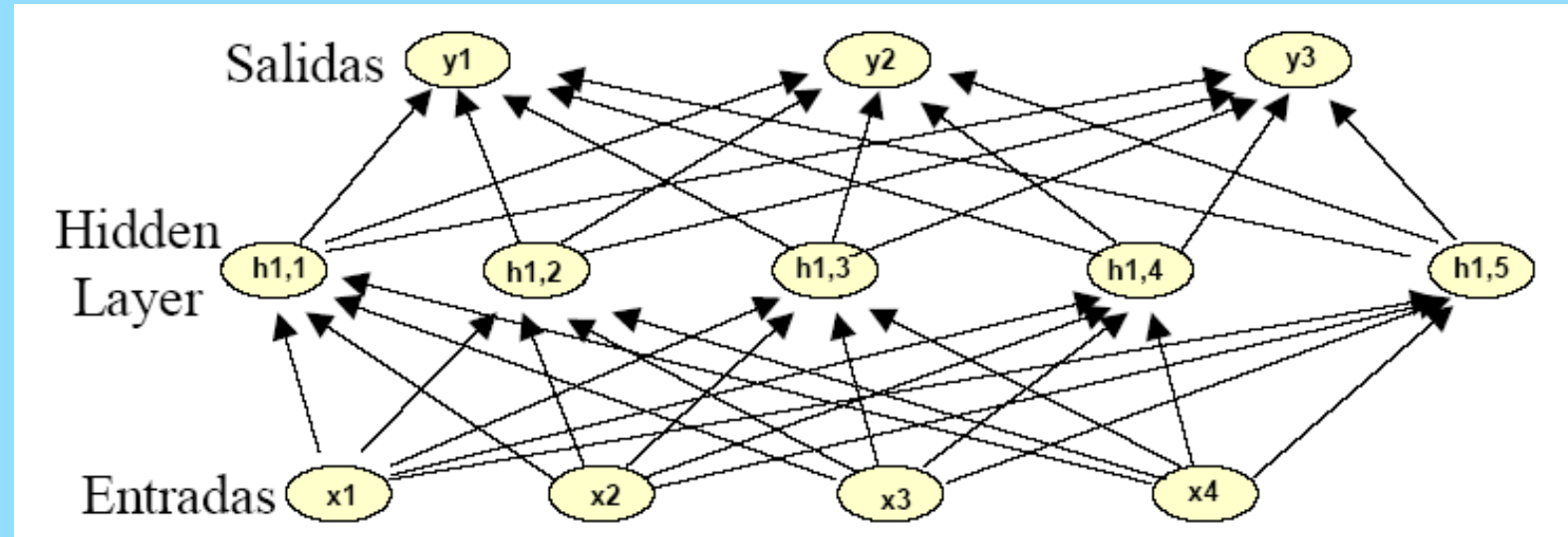
- ❖ COMPUTAN UNA FUNCIÓN LINEAL.
- ❖ PARA CADA y_j ES:

$$y'_j = \sum_{i=1}^n w_{i,j} \cdot x_i$$



DESCRIPCIÓN DE UNA RED NEURONAL (RN)

❑ MULTILAYER PERCEPTRON:



DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **FUNCIÓN DE SALIDA Y FUNCIONES DE TRANSFERENCIA O ACTIVACIÓN**
- ❑ **LAS CONEXIONES ENTRE ELEMENTOS DE PROCESO LLEVAN ASOCIADAS UN PESO O FUERZA DE CONEXIÓN w :**
 - ❖ **DETERMINA CUANTITATIVAMENTE EL EFECTO QUE PRODUCEN UNOS ELEMENTOS SOBRE OTROS.**
 - ❖ **EN LOS PESOS SE ALMACENA LA INFORMACIÓN DE LA RED.**
- ❑ **EL QUE UNA ENTRADA TENGA UN EFECTO EXCITATORIO O INHIBITORIO:**
 - ❖ **DEPENDE DE QUE EL SIGNO DEL PESO CORRESPONDIENTE SEA, RESPECTIVAMENTE, POSITIVO O NEGATIVO.**
- ❑ **LA EFECTIVIDAD DE LAS ENTRADAS ESTÁ DETERMINADA POR LA FUERZA DE LA CONEXIÓN:**
 - ❖ **REPRESENTADA POR EL VALOR ABSOLUTO DE LOS PESOS.**

DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **FUNCIÓN DE SALIDA Y FUNCIONES DE TRANSFERENCIA O ACTIVACIÓN**
- ❑ **C/U DE LOS ELEMENTOS w_{ij} DE LA MATRIZ DE PESOS w , CONOCIDA COMO PATRÓN DE CONEXIONES:**
 - ❖ **REPRESENTA LA INTENSIDAD Y SENTIDO DE LA RELACIÓN DEL ELEMENTO DE PROCESO j , CON RESPECTO AL ELEMENTO DE PROCESO i .**

DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **FUNCIÓN DE SALIDA Y FUNCIONES DE TRANSFERENCIA O ACTIVACIÓN**
- ❑ **EN UNA RN ARTIFICIAL ALIMENTADA HACIA DELANTE:**
 - ❖ **CON:**
 - **r ENTRADAS.**
 - **1 CAPA OCULTA CON q ELEMENTOS DE PROCESO.**
 - **1 UNIDAD DE SALIDA.**
 - ❖ **LA TRANSFORMACIÓN DE LAS ENTRADAS PUEDE RESUMIRSE EN LA FUNCIÓN DE SALIDA DE LA RED:**

- $$\hat{f}(x, W) = F\left(\beta_0 + \sum_{j=1}^q \beta_j G(x' \gamma_j)\right)$$

DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **FUNCIÓN DE SALIDA Y FUNCIONES DE TRANSFERENCIA O ACTIVACIÓN**

$$\hat{f}(x, W) = F\left(\beta_0 + \sum_{j=1}^q \beta_j G(x' \gamma_j)\right)$$

- **DONDE:**
- $\hat{f}(x, W)$ ES LA **SALIDA** DE LA RED.
- EL VECTOR $x = (1, x_1, x_2, \dots, x_r)'$ REPRESENTA LAS **ENTRADAS** DE LA RED.
- $\gamma_j = (\gamma_{j0}, \gamma_{j1}, \dots, \gamma_{ji}, \dots, \gamma_{jr})' \in \mathfrak{R}^{r+1}$ SON LOS **PESOS** DE LAS NEURONAS DE LA CAPA DE **ENTRADA** A LAS DE LA INTERMEDIA U **OCULTA**.
- $\beta_j, j = 0, \dots, q$, REPRESENTA LA **FUERZA DE CONEXIÓN** DE LAS UNIDADES **OCULTAS** A LAS DE **SALIDA**:
 - q ES EL NÚMERO DE **UNIDADES INTERMEDIAS**, ES DECIR, EL NÚMERO DE NODOS DE LA CAPA OCULTA.

DESCRIPCIÓN DE UNA RED NEURONAL (RN)

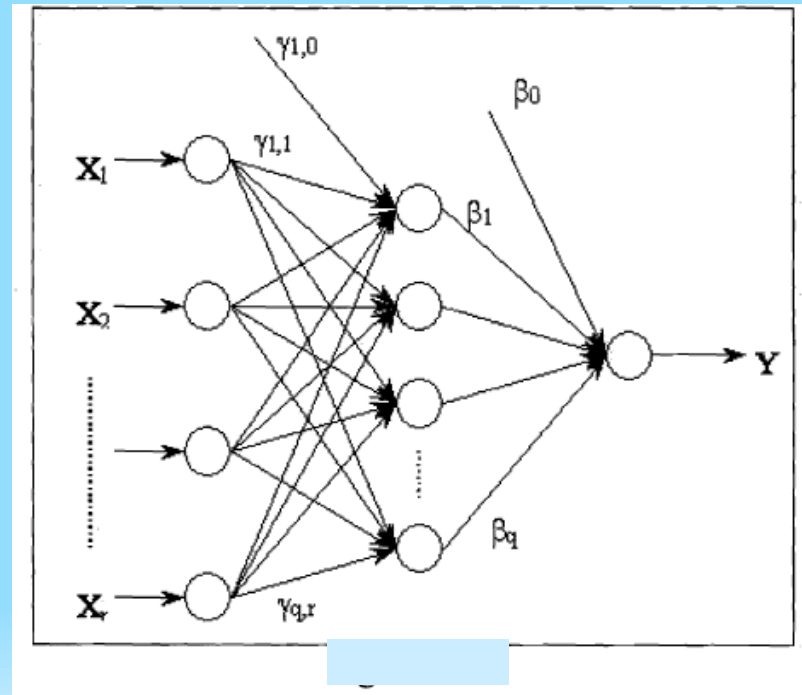
- ❑ **FUNCIÓN DE SALIDA Y FUNCIONES DE TRANSFERENCIA O ACTIVACIÓN**

$$\hat{f}(x, W) = F\left(\beta_0 + \sum_{j=1}^q \beta_j G(x' \gamma_j)\right)$$

- **$F: \mathfrak{R} \rightarrow \mathfrak{R}$ ES LA FUNCIÓN DE ACTIVACIÓN DE LA UNIDAD DE SALIDA.**
- **$G: \mathfrak{R} \rightarrow \mathfrak{R}$ ES LA FUNCIÓN DE ACTIVACIÓN DE LAS NEURONAS INTERMEDIAS.**
- **W ES UN VECTOR QUE INCLUYE TODOS LOS PESOS DE LA RED:**
 - γ_j
 - β_j .

DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **FUNCIÓN DE SALIDA Y FUNCIONES DE TRANSFERENCIA O ACTIVACIÓN**
 - ❖ FUNCIÓN $\hat{f}(x, W)$.



DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **FUNCIÓN DE SALIDA Y FUNCIONES DE TRANSFERENCIA O ACTIVACIÓN**
- ❑ **SE EMPLEAN COMO FUNCIONES DE ACTIVACIÓN FUNCIONES DE UMBRAL:**
 - ❖ **EL EFECTO ES QUE LAS UNIDADES SE ACTIVAN BRUSCAMENTE, ESTO ES, O NO SE ACTIVAN, O SE ACTIVAN DE GOLPE.**
 - ❖ **SON ADECUADAS PARA TAREAS DE CLASIFICACIÓN Y RECONOCIMIENTO.**
- ❑ **SE USAN TAMBIÉN FUNCIONES DE ACTIVACIÓN QUE PERMITEN QUE LAS NEURONAS SE ACTIVEN GRADUALMENTE A MEDIDA QUE EL NIVEL DE ACTIVIDAD DE SUS ENTRADAS AUMENTA:**
 - ❖ **UNA FUNCIÓN ES LA SIGMOIDAL O LOGÍSTICA:**
 - $G(a) = 1/(1 + \exp(-a))$:
 - **PRODUCE UNA RESPUESTA SIGMOIDAL ALISADA.**

DESCRIPCIÓN DE UNA RED NEURONAL (RN)

- ❑ **FUNCIÓN DE SALIDA Y FUNCIONES DE TRANSFERENCIA O ACTIVACIÓN**
- ❑ **OTRA POSIBILIDAD MUY USADA EN APLICACIONES ECONÓMICAS ES CONSIDERAR UNA RN:**
 - ❖ **ALIMENTADA HACIA DELANTE.**
 - ❖ **CON r ENTRADAS.**
 - ❖ **CON 1 CAPA OCULTA CON q ELEMENTOS DE PROCESO.**
 - ❖ **CON 1 UNIDAD DE SALIDA.**
 - ❖ **CON CONEXIONES DIRECTAS ENTRE LA CAPA DE ENTRADA Y LA DE SALIDA.**
 - ❖ **LA SALIDA DE LA RED SE OBTIENE MEDIANTE LA SIGUIENTE EXPRESIÓN:**

➤

$$\hat{f}(x, W) = x'\alpha + \beta_0 + \sum_{j=1}^q \beta_j G(x'\gamma_j)$$

DESCRIPCIÓN DE UNA RED NEURONAL (RN)

❑ **FUNCIÓN DE SALIDA Y FUNCIONES DE TRANSFERENCIA O ACTIVACIÓN**

➤

$$\hat{f}(x, W) = x'\alpha + \beta_0 + \sum_{j=1}^q \beta_j G(x'\gamma_j)$$

➤ α ES UN VECTOR DE DIMENSIÓN $r \times 1$ QUE REPRESENTA LOS PESOS DE LAS CONEXIONES DIRECTAS ENTRE LAS CAPAS DE ENTRADA Y SALIDA.

➤ W TIENE TODOS LOS PESOS DE LA RED:

- α
- γ_j
- β_j

APRENDIZAJE EN LAS REDES NEURONALES

APRENDIZAJE EN LAS RN

- ❑ DESPUÉS DE DISEÑAR UNA RN SE **PRETENDE** QUE PARA CIERTAS **ENTRADAS**, O PATRONES EJEMPLO SUMINISTRADOS A LA RED:
 - ❖ LA RN **GENERE UNA SALIDA DESEADA**.
- ❑ PARA ELLO SE **REQUIERE**:
 - ❖ QUE LA **TOPOLOGÍA** DE LA RED (ENTENDIDA COMO LA ESTRUCTURA DE LA RED) SEA **ADECUADA**.
 - ❖ QUE **APRENDA** A PROPORCIONAR SOLUCIONES CORRECTAS:
 - ES NECESARIO UN **PROCESO DE APRENDIZAJE O ENTRENAMIENTO**.
- ❑ EL **APRENDIZAJE** PUEDE ENTENDERSE COMO:
 - ❖ UN PROCEDIMIENTO DE **PRUEBA Y ERROR**.
 - ❖ PERMITE LA **ESTIMACIÓN ESTADÍSTICA** DE LOS PARÁMETROS DEL **MODELO** DE RED NEURONAL EMPLEADO.

APRENDIZAJE EN LAS RN

- ❑ **HAY TRES TIPOS BÁSICOS DE APRENDIZAJE.**
- ❑ **CUANDO EL ENTRENADOR PROPORCIONA A LA RED LA SALIDA DESEADA, SE DICE QUE EL APRENDIZAJE ES SUPERVISADO.**
- ❑ **EN CASO CONTRARIO NOS ENCONTRAMOS ANTE UN APRENDIZAJE NO SUPERVISADO.**
- ❑ **UN TIPO INTERMEDIO DE APRENDIZAJE ES EL REFORZADO O HÍBRIDO:**
 - ❖ **EL ENTRENADOR SÓLO PROPORCIONA A LA RED UNA INDICACIÓN DE SI LA RESPUESTA A UNA ENTRADA DADA ES BUENA O MALA.**
- ❑ **LAS RN CON APRENDIZAJE NO SUPERVISADO SON AQUELLAS QUE ENTRENAN SIN NECESIDAD DE UN SUPERVISOR O ENTRENADOR EXTERNO QUE PROPORCIONE A LA RED LA SALIDA DESEADA:**
 - ❖ **SON CAPACES DE ORGANIZAR SUS PARÁMETROS INTERNAMENTE ADAPTÁNDOSE AL ENTORNO DEL MEJOR MODO POSIBLE.**

APRENDIZAJE EN LAS RN

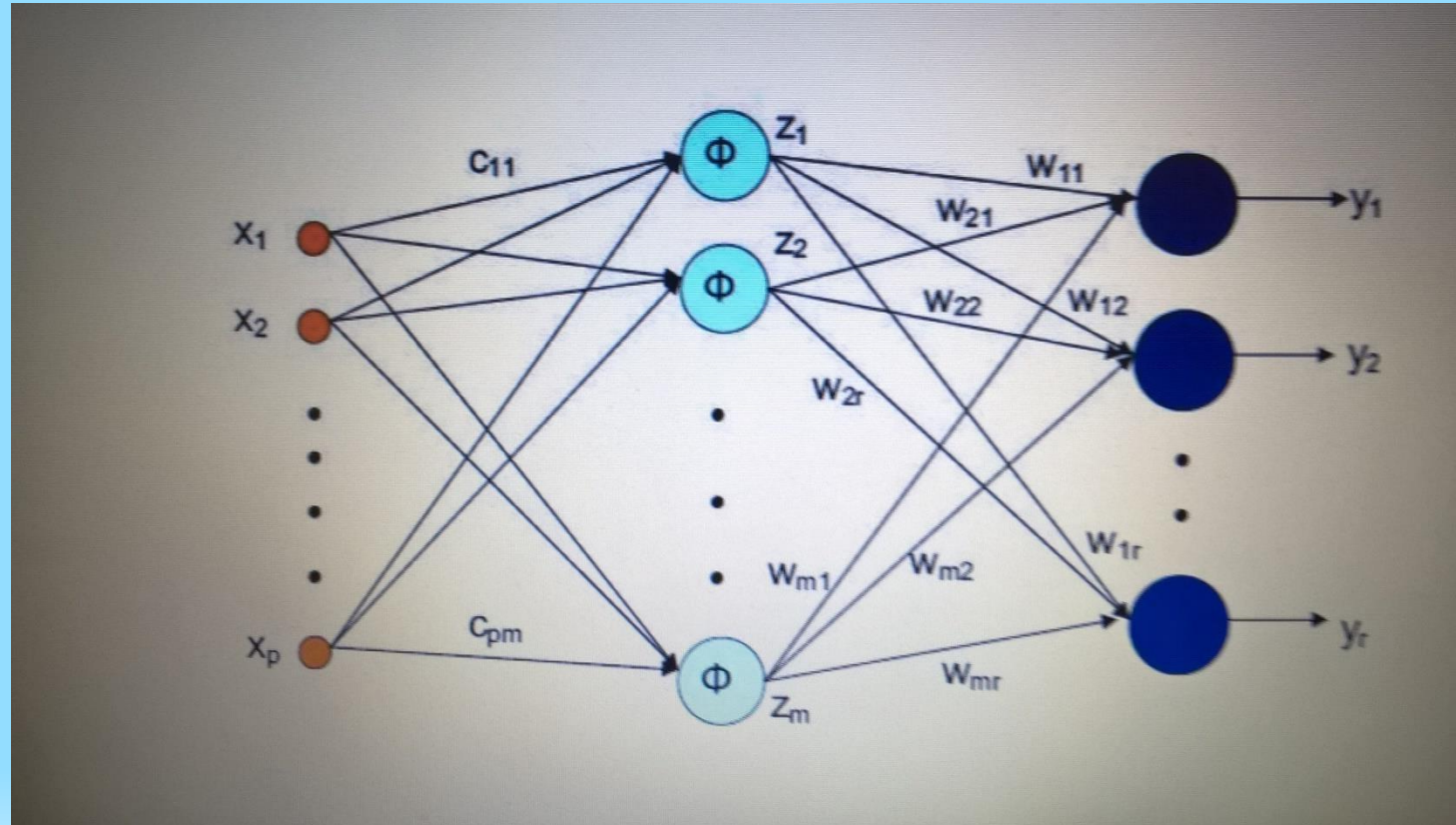
- ❑ LA **ESTRUCTURA** DE LA RED DEBE SER LA **ADECUADA** PARA EL TIPO DE DATOS DE **ENTRADA** QUE SE LE PRESENTAN.
- ❑ LOS SISTEMAS NEURONALES CON **APRENDIZAJE NO SUPERVISADO** SE CARACTERIZAN POR POSEER **ARQUITECTURAS SIMPLES**.
- ❑ LA MAYORÍA DE LAS RN SON **ALIMENTADAS HACIA ADELANTE**, O **FEED-FORWARD** CON UNA SOLA CAPA OCULTA.
- ❑ LOS **MODELOS MÁS CARACTERÍSTICOS** QUE ENTRENAN MEDIANTE **APRENDIZAJE NO SUPERVISADO** SON LAS **REDES DE KOHONEN Y DE GROSSBERG**.
- ❑ SUELEN UTILIZARSE PARA LA **CLASIFICACIÓN**.
- ❑ LAS REDES DE **KOHONEN** SUELEN UTILIZARSE:
 - ❖ CUANDO UNO DE LOS **OBJETIVOS** DEL ANÁLISIS SEA LA **VISUALIZACIÓN** SENCILLA E INTUITIVA DE LOS **CONGLOMERADOS**.
 - ❖ CUANDO SE **DESCONOCE** SU FORMA.
 - ❖ CUANDO EXISTAN CASOS **ATÍPICOS** O **ERRORES** EN LOS DATOS.

APRENDIZAJE EN LAS RN

- ❑ **LAS RN CON APRENDIZAJE SUPERVISADO:**
 - ❖ SE ASOCIAN AL **PERCEPTRÓN MULTICAPA (MULTILAYER PERCEPTRON: MLP)** Y LA **FUNCIÓN DE BASE RADIAL RBF**.
 - ❖ PRESENTAN UN PATRÓN DE **SALIDA** O VARIABLE DEPENDIENTE QUE LES PERMITE **CONTRASTAR Y CORREGIR** LOS DATOS.
- ❑ **LAS RN CON PATRÓN DE SALIDA SE UTILIZAN PARA LA CLASIFICACIÓN Y PARA LA PREDICCIÓN:**
 - ❖ SE PUEDEN SEGMENTAR MERCADOS, POSICIONAR PRODUCTOS, REALIZAR PREVISIONES DE DEMANDA, EVALUACIONES DE EXPEDIENTES DE CRÉDITO O DE ANÁLISIS DEL VALOR DE ACCIONES EN BOLSA Y MUCHAS APLICACIONES MÁS.

APRENDIZAJE EN LAS RN

- ❑ RED DE NEURONAS DE FUNCIÓN DE BASE RADIAL (RBF):



APRENDIZAJE EN LAS RN

- ❑ **EL APRENDIZAJE REFORZADO O HÍBRIDO ES INTERMEDIO ENTRE EL SUPERVISADO Y EL NO SUPERVISADO.**
- ❑ **EXISTE UN PROFESOR O SUPERVISOR EXTERNO.**
- ❑ **EL “ENTRENADOR” NO PROPORCIONA A LA RED LAS SALIDAS DESEADAS:**
 - ❖ **SU COMPORTAMIENTO SE EVALÚA DE MANERA GLOBAL.**
 - ❖ **SÓLO ES POSIBLE DECIDIR E INDICAR A LA RED SI SU RESPUESTA ES BUENA O MALA Y EN QUÉ GRADO SE COMPORTA BIEN.**
- ❑ **EL FUNDAMENTO DEL APRENDIZAJE REFORZADO ES QUE SE DEBEN REFORZAR AQUELLAS ACCIONES QUE GENERAN UNA MEJORA EN EL COMPORTAMIENTO Y RESPUESTA DE LA RED NEURONAL.**

APRENDIZAJE EN LAS RN

- ❑ FINALIZADO EL APRENDIZAJE SE DEBE PROCEDER A **TESTEAR LA RED.**
- ❑ CONSISTE EN **INTRODUCIR NUEVOS PATRONES DE ENTRADA Y COMPROBAR LA EFICACIA DEL SISTEMA GENERADO.**
- ❑ **SI NO RESULTA ACEPTABLE:**
 - ❖ **SE REPITE LA FASE DE ENTRENAMIENTO UTILIZANDO NUEVOS PATRONES-EJEMPLO.**
 - ❖ **INCLUSO PUEDE SER NECESARIO MODIFICAR LA ESTRUCTURA DE LA RED.**

FUNCIONAMIENTO DE UNA RED NEURONAL

FUNCIONAMIENTO DE UNA RN

- ❑ PARA LA **CREACIÓN Y APLICACIÓN** DE UNA RN A UN PROBLEMA CONCRETO SE DEBEN DISTINGUIR LOS SIGUIENTES PASOS:
- ❑ **CONCEPTUALIZACIÓN DEL MODELO** PARA EL ESTUDIO DEL PROBLEMA CONCRETO:
 - ❖ EN ESTE **MODELO** SE DEBE SEÑALAR LAS **ENTRADAS**, LAS **SALIDAS** Y LA **INFORMACIÓN** DE QUE SE DISPONE.
- ❑ **ADECUACIÓN DE LA INFORMACIÓN** DE QUE SE DISPONE A LA ESTRUCTURA DE LA RED A CREAR:
 - ❖ SE CONSTITUIRÁN:
 - **LOS PATRONES DE APRENDIZAJE:**
 - PARTE DE LA INFORMACIÓN QUE VA A SER UTILIZADA PARA EL **ENTRENAMIENTO** O **APRENDIZAJE** DE LA RED.
 - **LOS PATRONES DE VALIDACIÓN:**
 - PARTE DE LA INFORMACIÓN QUE VA A SER UTILIZADA COMO **VALIDACIÓN** DE LA RED.

FUNCIONAMIENTO DE UNA RN

❑ FASE DE APRENDIZAJE:

- ❖ SE LE VA PRESENTANDO A LA RED LOS PATRONES ADECUADOS.
- ❖ LA RED VA PROPORCIONANDO UNA SALIDA.
- ❖ ESTE PROCESO SE REPITE UN CIERTO N° DE ETAPAS.
- ❖ ESTAS SALIDAS SE COMPARAN CON LAS SALIDAS ESPERADAS.
- ❖ LOS ALGORITMOS DE APRENDIZAJE INTENTAN MINIMIZAR EL ERROR QUE HAY ENTRE LA SALIDA PROPORCIONADA POR LA RED Y LA SALIDA ESPERADA.

❑ FASE DE VALIDACIÓN:

- ❖ SE PRESENTAN A LA RED ENTRENADA EL CONJUNTO DE PATRONES DE VALIDACIÓN.
- ❖ SE VE EL ERROR COMETIDO POR LA RED EN ESTE CONJUNTO.
- ❖ ESTE ERROR ES UNA MEDIDA DE LA BONDAD DE LA RED.

FUNCIONAMIENTO DE UNA RN

- ❑ **FASE DE GENERALIZACIÓN:**
 - ❖ **SI SE HA CONSEGUIDO UNA RED ADECUADA SE PROCEDE A UTILIZAR LA RED COMO MODELO PREDICTOR.**
 - ❖ **APORTÁNDOLE UNA NUEVA ENTRADA, LA RED LA PROCESARÁ Y DARÁ UNA SALIDA.**

FUNCIONAMIENTO DE UNA RN

- ❑ **EJEMPLO:**
- ❑ SE CONSIDERA QUE UNA ENTIDAD BANCARIA QUIERE CONSTRUIR UN MODELO PARA LA CONCESIÓN DE CRÉDITOS PERSONALES PARA LA COMPRA DE UN AUTOMÓVIL.
- ❑ EL CONJUNTO DE INFORMACIÓN ESTÁ CONSTITUIDO POR UN FICHERO DE 5.000 CLIENTES ANTERIORES EN EL QUE CONSTA PARA EL CLIENTE LAS SIGUIENTES VARIABLES: SEXO, ESTADO CIVIL, SALARIO, BIENES, CUENTA CORRIENTE ETC., Y SI HA HECHO FRENTE A LOS PAGOS REGULARMENTE O NO.
- ❑ EL MODELO TENDRÍA:
 - ❖ UNA CAPA DE ENTRADA CON TANTOS NODOS COMO VARIABLES APARECEN EN EL FICHERO.
 - ❖ UNA O VARIAS CAPAS INTERMEDIAS.
 - ❖ UNA CAPA DE SALIDA CON UN NODO:
 - APARECE UNA SALIDA CERO O UNO SEGÚN HAYA HECHO FRENTE A LOS PAGOS REGULARMENTE O NO.

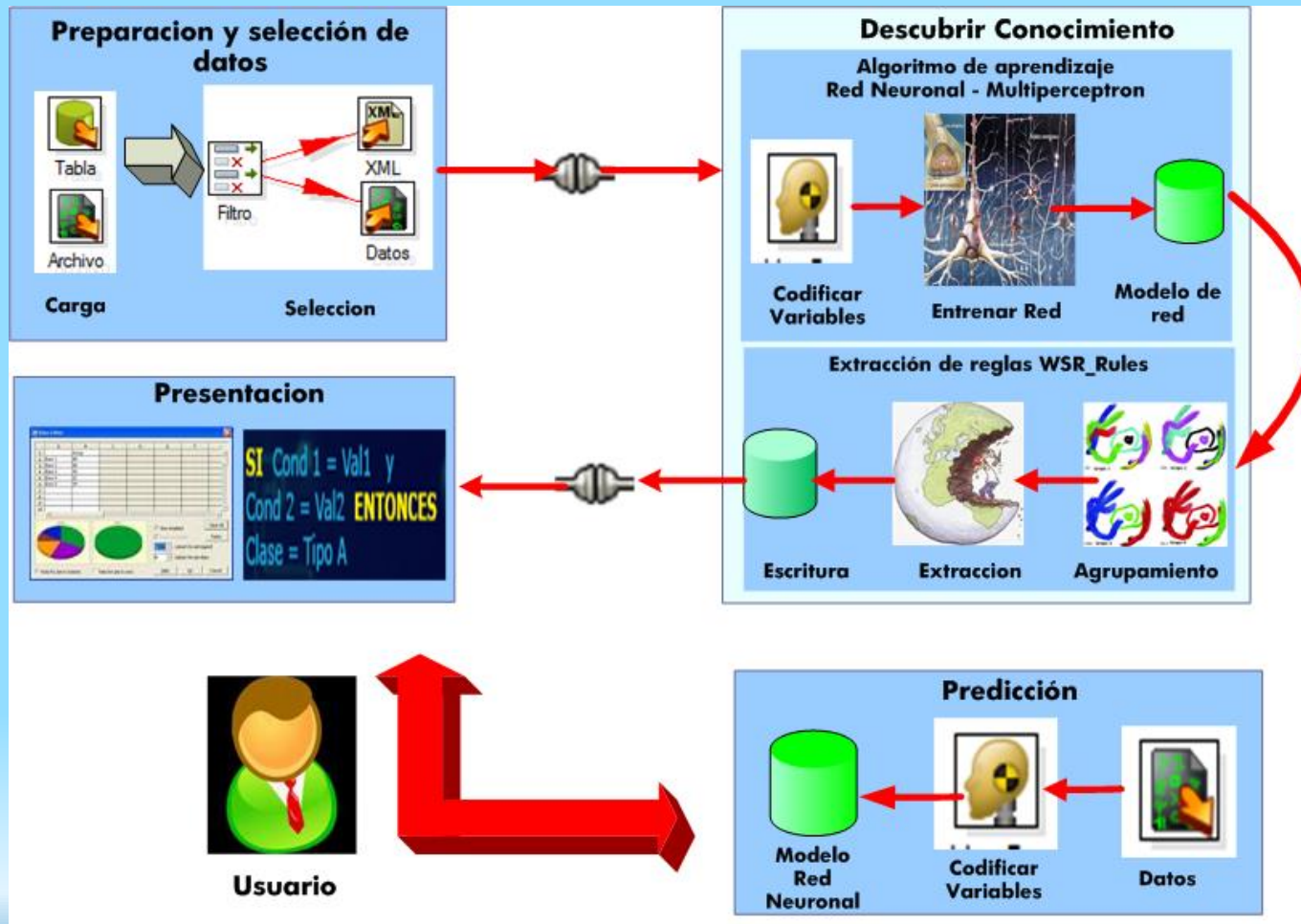
FUNCIONAMIENTO DE UNA RN

- ❑ **EJEMPLO:**
- ❑ **SE DIVIDE EL FICHERO EN UNA PARTE QUE SERÁ EL CONJUNTO DE APRENDIZAJE Y LA OTRA EL CONJUNTO DE VALIDACIÓN.**
- ❑ **A LA RED SE LE DA COMO ENTRADA UN CONJUNTO DE APRENDIZAJE, LA RED SE ENTRENARÁ.**
- ❑ **SE VALIDA EL MODELO CON EL CONJUNTO DE VALIDACIÓN.**
- ❑ **CUANDO LLEGUE EL NUEVO CLIENTE:**
 - ❖ **SE PRESENTA SU INFORMACIÓN A LA RN.**
 - ❖ **SI LA RN ESTÁ BIEN ENTRENADA CLASIFICARÁ AL CLIENTE:**
 - **EN LA CLASE DE LOS QUE PAGA EN TIEMPO Y FORMA, O.**
 - **EN LA CLASE COMPLEMENTARIA.**

FUNCIONAMIENTO DE UNA RN

❑ EJEMPLO:

❖ DETECCIÓN DE FRAUDE.



ANÁLISIS DISCRIMINANTE A TRAVÉS DEL PERCEPTRÓN

ANÁLISIS DISCRIMINANTE A TRAVÉS DEL PERCEPTRÓN

- ❑ EL PROPÓSITO DEL PERCEPTRÓN ES CLASIFICAR UN CONJUNTO DE SEÑALES DE ENTRADA (ESTÍMULOS O PATRONES DE ENTRADA), EN UNA DE DOS CLASES, C_1 o C_2 .
- ❑ LA REGLA DE DECISIÓN PARA LA CLASIFICACIÓN CONSISTE EN ASIGNAR LAS ENTRADAS x_1, x_2, \dots, x_n CORRESPONDIENTES A UN PATRÓN DADO:
 - ❖ A LA CLASE C_1 SI LA SALIDA DEL PERCEPTRÓN ES +1.
 - ❖ A LA CLASE C_2 SI LA SALIDA DEL PERCEPTRÓN ES -1.
- ❑ LA FUNCIÓN QUE EXPRESA LA SALIDA, y , DEL PERCEPTRÓN COMO UNA FUNCIÓN DE LAS ENTRADAS, x_1, x_2, \dots, x_n ES:
 - ❖ $y = \varphi_2(u - \theta)$,
 - ❖ $u = \sum_{j=1}^n w_j x_j$ ES EL POTENCIAL SINÁPTICO.
 - ❖ $\theta \in \mathfrak{R}$ ES EL UMBRAL.
 - ❖ $y = \begin{cases} 1 & \text{si } u \geq \theta \\ -1 & \text{si } u < \theta \end{cases}$

ANÁLISIS DISCRIMINANTE A TRAVÉS DEL PERCEPTRÓN

- ❑ LA PRIMERA CUESTIÓN ES **CÓMO DETERMINAR LOS PESOS SINÁPTICOS** DE MANERA QUE EL PERCEPTRÓN CONSIGA UNA **CLASIFICACIÓN CORRECTA** DE CADA UNO DE LOS PATRONES DE ENTRADA.
- ❑ SE SUPONE QUE SE TIENE p **PATRONES DE ENTRADA**, REPRESENTADOS POR $(x_1^r, x_2^r, \dots, x_n^r)$, $r = 1, 2, \dots, p$.
- ❑ SE CONOCE LA CLASE A LA QUE PERTENECE C/U DE ELLOS (LOS VECTORES O PATRONES DE ENTRENAMIENTO).
- ❑ **EJEMPLO:**
 - ❖ $(x_1^r, x_2^r, \dots, x_n^r)$ PUEDE REPRESENTAR UN CONJUNTO DE n **SÍNTOMAS DEL PACIENTE** r QUE PUEDE ($y=1$), O NO ($y=-1$), PRESENTAR UNA CIERTA ENFERMEDAD.

ANÁLISIS DISCRIMINANTE A TRAVÉS DEL PERCEPTRÓN

- ES POSIBLE MODIFICAR LOS PESOS SINÁPTICOS A TRAVÉS DE UN PROCEDIMIENTO ITERATIVO, EL PROCESO DE APRENDIZAJE DE LA RED:
 - ❖ DESPUÉS DE UN N° FINITO DE ITERACIONES DEBE CLASIFICAR CORRECTAMENTE LOS p PATRONES DE ENTRADA, SIEMPRE QUE SEA POSIBLE:
 - LAS CLASES C_1 y C_2 DEBEN SER SEPARABLES LINEALMENTE.
 - DOS CLASES SE DICE QUE SON LINEALMENTE SEPARABLES SI EXISTE UN CONJUNTO DE PESOS, w_1, w_2, \dots, w_n , TAL QUE:
 - $\sum_{j=1}^n w_j x_j > \theta, \forall (x_1, x_2, \dots, x_n) \in C_1$
 - $\sum_{j=1}^n w_j x_j < \theta, \forall (x_1, x_2, \dots, x_n) \in C_2$

ANÁLISIS DISCRIMINANTE A TRAVÉS DEL PERCEPTRÓN

- ❑ **EL PROCESO DE APRENDIZAJE CONSISTE EN OBSERVAR LOS VALORES DE SALIDA PARA C/U DE LOS PATRONES DE ENTRENAMIENTO Y COMPROBAR SI ES CORRECTA LA CLASIFICACIÓN REALIZADA POR EL PERCEPTRÓN.**
- ❑ **EN CASO DE CLASIFICACIÓN INCORRECTA SE MODIFICAN LOS PESOS.**
- ❑ **SE REPITE EL PROCESO, PATRÓN A PATRÓN, HASTA CONSEGUIR:**
 - ❖ **CLASIFICAR CORRECTAMENTE A TODOS LOS PATRONES DE ENTRENAMIENTO, O.**
 - ❖ **QUE EL N° DE CLASIFICACIONES INCORRECTAS SE CONSIDERE ADECUADO.**

ANÁLISIS DE SERIES TEMPORALES MEDIANTE REDES NEURONALES

ANÁLISIS DE SERIES TEMPORALES

MEDIANTE REDES NEURONALES

- ❑ UNA SERIE TEMPORAL CONSISTE EN UNA SECUENCIA DE VALORES DE VARIAS VARIABLES QUE EVOLUCIONAN (VAN CAMBIANDO) EN EL TIEMPO.
- ❑ SE TRATA DE PREDECIR EL COMPORTAMIENTO FUTURO DEL FENÓMENO O SISTEMA DINÁMICO QUE GENERA ESOS VALORES BASÁNDOSE EN UNA COLECCIÓN DE DATOS HISTÓRICOS.
- ❑ EJEMPLO: LA PREDICCIÓN DEL CONSUMO DE ENERGÍA ELÉCTRICA O LA PREDICCIÓN DEL N° DE VACUNAS CONTRA LA GRIPE QUE SE VAN A DEMANDAR EN UNA REGIÓN DETERMINADA.
- ❑ LA MEJOR MANERA DE RESOLVER ESTO ES ENCONTRANDO LA LEY SUBYACENTE QUE GENERA DICHS PROCESOS:
 - ❖ SE PUEDE OBTENER MEDIANTE MÉTODOS ANALÍTICOS, COMO PUEDE SER UN CONJUNTO DE ECUACIONES DIFERENCIALES.
- ❑ LA INFORMACIÓN QUE SE TENDRÁ DEL PROCESO VA A SER GENERALMENTE PARCIAL O INCOMPLETA:
 - ❖ NO SE DISPONE DE UN MODELO ANALÍTICO CONOCIDO.

ANÁLISIS DE SERIES TEMPORALES

MEDIANTE REDES NEURONALES

- ❑ **SE INTENTARÁ DESCUBRIR ALGUNA REGULARIDAD EMPÍRICA FUERTE EN LAS OBSERVACIONES DE LAS SERIES TEMPORALES:**
 - ❖ **EN MUCHOS PROBLEMAS DEL MUNDO REAL ALGUNAS REGULARIDADES, COMO LA PERIODICIDAD, APARECEN ENMASCARADAS POR RUIDOS.**
 - ❖ **ALGUNOS PROCESOS DINÁMICOS SE DESCRIBEN POR SERIES DE TIEMPO CAÓTICAS, DONDE LOS DATOS PARECEN ALEATORIOS SIN PERIODICIDADES APARENTES.**
 - ❖ **EL CAOS IMPIDE CUALQUIER PREDICCIÓN A LARGO PLAZO.**
 - ❖ **SE CONSIGUEN BUENOS RESULTADOS PARA PREDICCIONES A CORTO PLAZO UTILIZANDO RN QUE EMPLEAN UNA FUNCIÓN MULTIVARIANTE NO LINEAL:**
 - $y(t)=F[y(t-1), y(t-2), \dots, y(t-k)]$
 - $y(t), t = m, m-1, \dots, k$, SON LAS MUESTRAS DADAS DE LA SERIE DE TIEMPO.
 - F ES UNA FUNCIÓN NO LINEAL DESCONOCIDA.
 - $k < m$.

ANÁLISIS DE SERIES TEMPORALES

MEDIANTE REDES NEURONALES

- ❑ LA MAYORÍA DE LAS TÉCNICAS DISPONIBLES SUPONEN **RELACIONES LINEALES:**
 - ❖ ENTRE LAS VARIABLES, O.
 - ❖ ENTRE LAS VARIABLES DESFASADAS EN EL TIEMPO.
- ❑ LA FORMULACIÓN DE **MODELOS NO LINEALES RAZONABLES** ES UNA TAREA **MUY DIFÍCIL**.
- ❑ LAS **RN** SON UNA BUENA **ALTERNATIVA** PARA EL CÁLCULO DE **PREDICCIONES**.
- ❑ EL **PROBLEMA PRINCIPAL** ES LA DIFICULTAD DE ENCONTRAR LA **RED ADECUADA PARA CADA CASO**.

ANÁLISIS DE SERIES TEMPORALES

MEDIANTE REDES NEURONALES

- ❑ EL PERCEPTRÓN MULTICAPA ES EL MODELO MÁS UTILIZADO DE RN PARA PREDICCIÓN DE VALORES FUTUROS.
- ❑ LA UNIDAD DE SALIDA DA UNA COMBINACIÓN LINEAL DE LAS SALIDAS DE TODAS LAS UNIDADES OCULTAS:

$$\diamond \hat{y}(t) = \omega_0 + \sum_{j=1}^h \omega_j \Psi_j \left(\sum_{i=1}^k \omega_{ji} y(t-i) + \omega_{j0} \right)$$

- ❖ Ψ_j ES LA FUNCIÓN DE ACTIVACIÓN.
 - ❖ LOS PESOS SINÁPTICOS ω_{ji} Y ω_j SE VAN AJUSTANDO DURANTE EL PROCESO DE APRENDIZAJE Y PUEDEN SER POSITIVOS, NEGATIVOS O NULOS.
- ❑ PARA DISEÑAR LA RED NEURONAL SE TIENE QUE VER CUÁNTAS NEURONAS OCULTAS DEBEMOS UTILIZAR:
 - ❖ GENERALMENTE ES PROPORCIONAL AL TAMAÑO DE LA MUESTRA QUE SE UTILIZA PARA EL ENTRENAMIENTO DE LA RED.

ANÁLISIS DE SERIES TEMPORALES

MEDIANTE REDES NEURONALES

- ❑ EL COMPORTAMIENTO DE LA RED SE VALORA SEGÚN LA **FUNCIÓN DE ERROR**:
 - ❖ $E(\omega) = \sum_{k=1}^p (y(k) - \hat{y}(k))^2$
 - ❖ p ES EL N° DE **MUESTRAS** UTILIZADAS EN EL **ENTRENAMIENTO**.
- ❑ SE HA DEMOSTRADO QUE PARA p MUESTRAS DE **ENTRENAMIENTO**:
 - ❖ UN PERCEPTRÓN CON UNA **SÓLA CAPA OCULTA** DE $p-1$ NEURONAS PUEDE **IMPLEMENTAR** DICHO CONJUNTO DE ENTRENAMIENTO.
 - ❖ SON **SUFICIENTES** $p-1$ NEURONAS, PERO GENERALMENTE SE NECESITARÁN MENOS.

ANÁLISIS DE COMPONENTES PRINCIPALES CON REDES NEURONALES

ANÁLISIS DE COMPONENTES PRINCIPALES CON RN

- ❑ SE SABE QUE LA OBTENCIÓN DE LAS COMPONENTES PRINCIPALES PERMITIRÁ:
 - ❖ UNA DISMINUCIÓN DE LA DIMENSIONALIDAD PERDIENDO POCA INFORMACIÓN.
 - ❖ LA DETECCIÓN DE LOS RASGOS CARACTERÍSTICO DEL CONJUNTO DE DATOS.
- ❑ PUEDE SER ÚTIL EN PROBLEMAS DE CLASIFICACIÓN Y RECONOCIMIENTO.
- ❑ SE CONSIDERARÁ UNA RN CUYA SALIDA VIENE DADA POR:
 - ❖ $y = \omega'x$.
 - ❖ $x = (x_1, x_2, \dots, x_n)'$ EL VECTOR DE ENTRADA A LA UNIDAD DE PROCESO (NEURONA).
 - ❖ $\omega = (\omega_1, \omega_2, \dots, \omega_n)'$ EL VECTOR DE PESOS SINÁPTICOS.

ANÁLISIS DE COMPONENTES PRINCIPALES CON RN

- ❑ LA REGLA DE APRENDIZAJE QUE SE CONSIDERARÁ FUE PROPUESTA POR HEBB (APRENDIZAJE HEBBIANO).
- ❑ EL POSTULADO DE HEBB ESTABLECE QUE LAS VARIACIONES DE LOS PESOS SINÁPTICOS (CAMBIOS DE LA EFICACIA SINÁPTICA) SON PROPORCIONALES AL PRODUCTO DE:
 - ❖ LAS SEÑALES PRESINÁPTICAS (ENTRADA), x .
 - ❖ LA SEÑAL POSTSINÁPTICA, y .
- ❑ ES DECIR:
 - ❖ $\Delta\omega_j(k+1) = \omega_j(k+1) - \omega_j(k) = \eta x_j(k)y(k), j = 1, 2, \dots, n$
- ❑ EN FORMA VECTORIAL:
 - ❖ $\Delta\omega(k+1) = \eta y(k) x(k)$, SIENDO η LA TASA DE APRENDIZAJE.

ANÁLISIS DE COMPONENTES

PRINCIPALES CON RN

- ❑ SE SUPONE AHORA QUE EL VECTOR DE ENTRADAS x ES UN VECTOR ALEATORIO CON FUNCIÓN DE DENSIDAD $f(x)$.
- ❑ EN CADA INSTANTE DE TIEMPO k , $k = 1, 2, \dots$ SE LE PRESENTA A LA RED UN VECTOR DE ENTRADA x .
- ❑ ES UN VALOR DEL VECTOR ALEATORIO QUE TIENEN COMO FUNCIÓN DE DENSIDAD $f(x)$.
- ❑ SE VAN ACTUALIZANDO LOS PESOS SINÁPTICOS SEGÚN LA REGLA DE HEBB.
- ❑ SE PUEDE CONSIDERAR QUE:
 - ❖ $\Delta\omega(k+1) =$ ES UN VECTOR ALEATORIO CUYO VALOR ESPERADO ES:
 - $E[\Delta\omega(k+1)] = E[\eta x(k)y(k)] = E[\eta x(k)(x(k)'\omega(k))] = \eta E[(x(k)x(k)')\omega(k)]$
 - ❖ SI SE SUPONE QUE x Y ω SON ESTADÍSTICAMENTE INDEPENDIENTES, LA EXPRESIÓN ANTERIOR ES:
 - $\eta E[x(k)x(k)']E[\omega(k)]$

CLUSTERING MEDIANTE REDES NEURONALES

CLUSTERING MEDIANTE RN

- ❑ LAS TÉCNICAS DE **CLUSTERING** SE OCUPAN DE REALIZAR UNA **PARTICIÓN** DE UN CONJUNTO DADO DE **DATOS** EN VARIOS **GRUPOS** O **CATEGORÍAS**.
- ❑ TIENEN **APLICACIONES** EN MUCHOS CAMPOS COMO LA **MEDICINA**, LA **BIOLOGÍA**, LA **ZOOLOGÍA** Y, SOBRE TODO, EN LAS **TECNOLOGÍAS** DE LA **INFORMACIÓN**.
- ❑ LA MAYORÍA DE LAS **TÉCNICAS** QUE SE HAN DESARROLLADO SE PUEDE DIVIDIR EN TRES GRANDES GRUPOS:
 - ❖ **JERÁRQUICAS**.
 - ❖ **DE MINIMIZACIÓN DE COSTES** O **DISTORSIONES**.
 - ❖ **DE TEORÍA DE GRAFOS**.
- ❑ CUANDO SE TRATA DE **GRANDES CONJUNTOS** DE **DATOS**, SÓLO LAS **TÉCNICAS** DE **MINIMIZACIÓN DE COSTES** SON LAS **ADECUADAS**, POR RAZONES DE **COMPLEJIDAD COMPUTACIONAL**.

CLUSTERING MEDIANTE RN

- ❑ SE TIENE UN **CONJUNTO DE DATOS** $\{x_i \in R^n, i = 1, 2, \dots, p\}$.
- ❑ LOS DATOS CORRESPONDEN A n **VALORES DE UN VECTOR ALEATORIO** X CUYA **FUNCIÓN DE DISTRIBUCIÓN** F_x SE DESCONOCE:
 - ❖ SE SABE QUE ES UNA MIXTURA DE m FUNCIONES DE DISTRIBUCIÓN.
- ❑ EL PROBLEMA CONSISTE EN FORMAR m **GRUPOS O CATEGORÍAS** CON LOS DATOS SEGÚN UNA CIERTA FUNCIÓN DE **DISTANCIA** O MEDIDA DE DISTORSIÓN:
 - ❖ EN ESTE CASO SERÁ EL **ERROR CUADRÁTICO** (LA DISTANCIA EUCLÍDEA AL CUADRADO).

CLUSTERING MEDIANTE RN

- ❑ CADA **GRUPO** ESTARÁ **REPRESENTADO** POR SU **CENTROIDE**, ES DECIR, SU VECTOR MEDIA.
- ❑ SE TRATA DE **DETERMINAR** m **CENTROIDES**, $\omega_i \in R^n, i = 1, 2, \dots, m$, DE MANERA QUE LA **DISTORSIÓN** ENTRE CADA DATO Y SU CENTROIDE SEA MÍNIMA, ES DECIR, **MINIMIZAR**:
 - ❖ $J(\omega_1, \omega_2, \dots, \omega_m) = \sum_{i=1}^n \sum_{x_j \in C_i} \|x_j - \omega_i\|^2$
- ❑ LA **SOLUCIÓN** DE ESTE PROBLEMA CONDUCE A UNA **PARTICIÓN** DEL CONJUNTO DE DATOS EN m **CLASES**, CATEGORÍAS O REGIONES, $C_i = 1, 2, \dots, m$:
 - ❖ $x_j \in C_i \leftrightarrow \|x_j - \omega_i\| < \|x_j - \omega_k\|, \forall k \neq i$

CLUSTERING MEDIANTE RN

- ❑ PARA **RESOLVER** ESTAS AGRUPACIONES SE PUEDEN UTILIZAR LAS **RN COMPETITIVAS CON APRENDIZAJE NO SUPERVISADO**:
 - ❖ PERMITEN FORMAR **GRUPOS O CATEGORÍAS** CON LOS DATOS DE ENTRADA.
 - ❖ CADA ENTRADA SE **ASIGNARÁ** A LA **CLASE** QUE LE ESPECIFICA LA UNIDAD DE **SALIDA** ACTIVADA.
 - ❖ LAS NEURONAS **COMPITEN** ENTRE ELLAS PARA VER CUÁL ES LA QUE SE **ACTIVA**:
 - **SÓLO UNA** DE ELLAS SE ACTIVA.
 - ❖ AQUÍ LA RN, POR SÍ SOLA (**SIN SUPERVISIÓN**):
 - REALIZA LA **AGRUPACIÓN** DE LOS DATOS EN **CATEGORÍAS**.
 - DA LOS **PROTOTIPOS** DE CADA CATEGORÍA.

CLUSTERING MEDIANTE RN

- ❑ ESTE PARADIGMA DE APRENDIZAJE FUE INTRODUCIDO POR RUMELHART Y ZIPSER (1985).
- ❑ LA RN CONSTA DE UNA CAPA DE m NEURONAS O UNIDADES DE PROCESO QUE ESTÁN CONECTADAS CON LAS SEÑALES O PATRONES DE ENTRADA.
- ❑ DICHAS CONEXIONES TIENEN ASOCIADOS UNOS PESOS SINÁPTICOS QUE VAN A REPRESENTAR A LOS PATRONES PROTOTIPO.
- ❑ CADA NEURONA SÓLO TIENE DOS POSIBLES VALORES DE SALIDA:
 - ❖ 0 (NO ACTIVADA).
 - ❖ 1 (ACTIVADA).
 - ❖ SÓLO SE ACTIVA UNA NEURONA PARA CADA PATRÓN DE ENTRADA, AQUELLA QUE RECIBE UN MAYOR POTENCIAL SINÁPTICO.

CLUSTERING MEDIANTE RN

- SI LA NEURONA i :
 - ❖ RECIBE LA SEÑAL DE ENTRADA $x = (x_1, x_2, \dots, x_n)'$.
 - ❖ TIENE LOS PESOS SINÁPTICOS $\omega_i = (\omega_{i1}, \omega_{i2}, \dots, \omega_{in})'$.
 - ❖ SU POTENCIAL SINÁPTICO VIENE DADO POR LA EXPRESIÓN:
 - $h_i = \sum_{j=1}^n \omega_{ij}x_j - \frac{1}{2} \sum_{j=1}^n \omega_{ij}^2$
- CUANDO $h_r = \max \{h_1, h_2, \dots, h_m\}$ LA SALIDA DE LA RED ES:
 - ❖ $(y_1, y_2, \dots, y_{r-1}, y_{r+1}, \dots, y_m) = (0, 0, \dots, 0, 1, 0, \dots, 0)$
 - ❖ $h_i < h_j \leftrightarrow d(\omega_i, x) > d(\omega_j, x)$
 - ❖ SIENDO d LA DISTANCIA EUCLÍDEA.
- LA NEURONA QUE SE **ACTIVA** ES AQUELLA CUYO VECTOR DE **PESOS SINÁPTICOS ESTÁ MÁS PRÓXIMO** AL PATRÓN DE **ENTRADA** x :
 - ❖ ES EL MÁS PARECIDO.

CLUSTERING MEDIANTE RN

- ❑ SE UTILIZA LA REGLA DE APRENDIZAJE COMPETITIVO.
- ❑ PARA DEDUCIR ESTA REGLA:
 - ❖ SE CONSIDERA UN CONJUNTO DE p PATRONES DE ENTRADA:
 - $x_i = (x_{i1}, x_{i2}, \dots, x_{in})'$ $i = 1, 2, \dots, p$, CONJUNTO DE ENTRENAMIENTO DE LA RED.
 - ❖ A PARTIR DE ESTOS VALORES SE OBTENDRÁ EL VALOR DE LOS PESOS SINÁPTICOS DE CADA NEURONA DE LA RED.
 - ❖ SI PARA LA ENTRADA x RESULTA GANADORA LA NEURONA r ES PORQUE SU VECTOR DE PESOS ES EL MÁS PARECIDO Y ASÍ ESTA ENTRADA SE ASIGNA A LA CLASE r .

CLUSTERING MEDIANTE RN

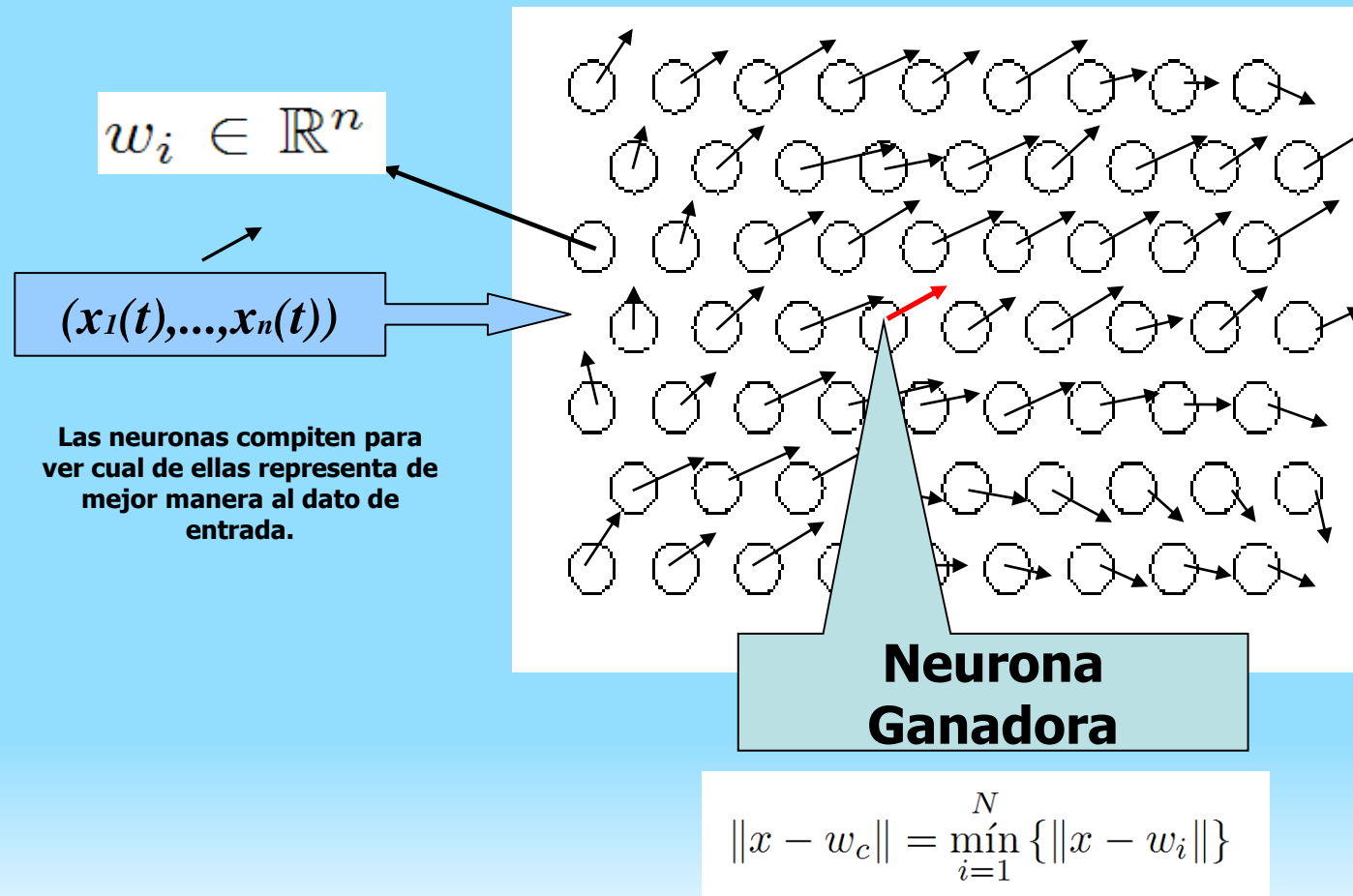
- ❑ ESTO SIGNIFICA QUE EL VECTOR DE PESOS SINÁPTICOS ω_r ES EL QUE REPRESENTA AL PATRÓN DE ENTRADA x .
- ❑ PARA QUE ESTA REPRESENTACIÓN SEA LO MEJOR POSIBLE:
 - ❖ ω_r DEBE ESTAR LO MÁS PRÓXIMO POSIBLE A TODOS LOS PATRONES DEL GRUPO QUE REPRESENTA.
 - ❖ SE DEBE ACTUALIZAR (MODIFICAR) EL VECTOR DE PESOS SINÁPTICOS ω_r DE LA NEURONA GANADORA r ACERCÁNDOLO AL PATRÓN DE ENTRADA s .
 - ❖ SI EN LA ETAPA $k+1$ SE HA INTRODUCIDO EL PATRÓN DE ENTRADA x Y HA RESULTADO GANADORA LA NEURONA r ENTONCES SE MODIFICAN SUS PESOS SINÁPTICOS SEGÚN LA EXPRESIÓN:
 - $\omega_r(k+1) = (1 - \eta_r(k)) \omega_r(k) + \eta_r(k)x(k) = \omega_r(k) + \eta_r(k)[x(k)\omega_r(k)]$
 - η_r ES LA TASA DE APRENDIZAJE DE LA NEURONA r .

CLUSTERING MEDIANTE RN

- ❑ LAS NEURONAS APRENDEN INDIVIDUALMENTE A ESPECIALIZARSE SOBRE CONJUNTOS DE PATRONES Y SON ASÍ DETECTORAS DE CARACTERÍSTICAS.
- ❑ CADA VECTOR DE PESOS SINÁPTICOS REPRESENTA EL PROTOTIPO DE LA CLASE O GRUPO QUE DEFINE LA NEURONA CORRESPONDIENTE.
- ❑ AHALT Y OTROS (1990) HAN DESTACADO LA CONVENIENCIA DE UTILIZAR EL APRENDIZAJE COMPETITIVO PORQUE CONDUCE A SOLUCIONES ÓPTIMAS O CERCANAS A LA ÓPTIMA.

CLUSTERING MEDIANTE RN

□ APRENDIZAJE COMPETITIVO



MUCHAS GRACIAS

!!! ???