

# Las Matemáticas del Aprendizaje Automático Conexionista y su relación con las TIC's

**Dra. M. Pilar Gómez Gil**

Coordinación de Computación  
Instituto Nacional de Astrofísica, Óptica y Electrónica

[pgomez@inaoep.mx](mailto:pgomez@inaoep.mx)

Oficina 8228



[ccc.inaoep.mx/~pgomez](http://ccc.inaoep.mx/~pgomez)

# Algo sobre mi lugar de trabajo

- El Instituto Nacional de Astrofísica, Óptica y Electrónica es uno de los 27 “centros Conacyt”
- Su misión es:  
*“contribuir a la generación, avance y difusión del conocimiento para el desarrollo del país y de la humanidad, por medio de la identificación y solución de problemas científicos y tecnológicos y de la formación de especialistas en las áreas de Astrofísica, Óptica, Electrónica, Ciencias Computacionales y áreas afines”*

# El Gran Telescopio Milimétrico, un proyecto del INAOE

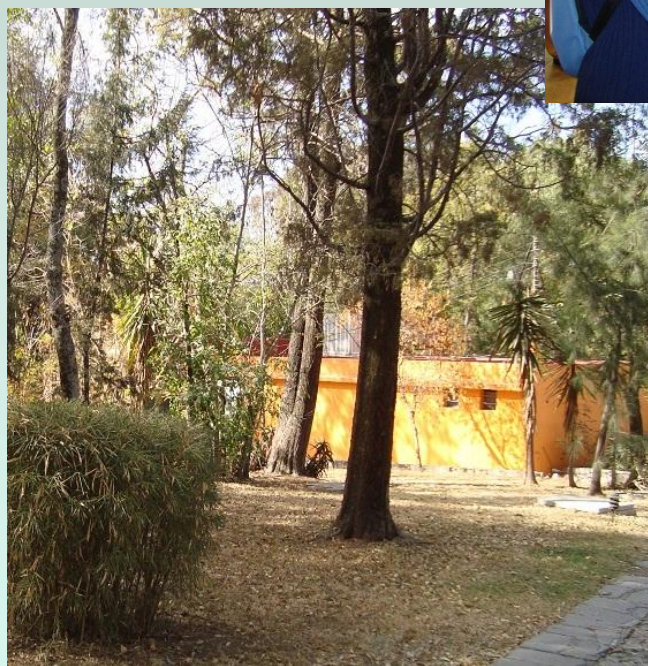


<http://www.lmtgim.org/images/sitepics/inauguration/GTM-22-11-06.jpg>

# La coordinación de Computación

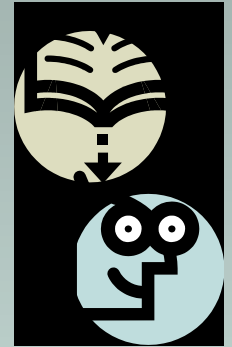
- Formada por 21 investigadores, 2 investigadores visitantes, 2 técnicos en investigación, 40 alumnos de maestría y 30 de doctorado
- LÍNEAS Y LABORATORIOS DE investigación.
  - > Aprendizaje Computacional y Reconocimiento de Patrones
  - > Cómputo Reconfigurable y de Alto Rendimiento
  - > Cómputo y Procesamiento Ubicuo
  - > Procesamiento de Bioseñales y Computación Médica
  - > Robótica
  - > Tecnologías del Lenguaje
  - > Visión por Computadora



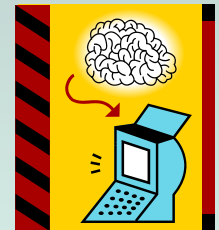


(c) P. GOMEZ, INAOE 2013

# Laboratorio "Machine Learning and Pattern Recognition" - MLPR (Aprendizaje Computacional y Reconocimiento de Patrones)



- Enseñar a las computadoras a realizar tareas complejas que serían difíciles de programar directamente
- Identificar información y/o conocimiento útil a partir de datos



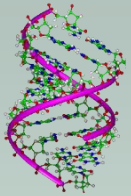
# Objetivos del MLPR

- Analizar grandes cantidades de datos
- Extraer y utilizar automáticamente conocimiento
- Facilitar la programación
- Proveer capacidad de adaptación
- Aprender tareas difíciles





# Ejemplos de Aplicaciones



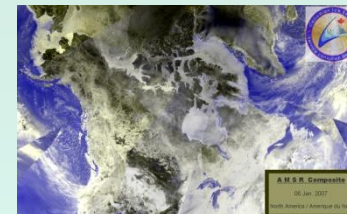
- Biología: entender los procesos que dan origen a la vida

- Banca: detección automática de fraudes, asignación de calificaciones crediticias y pronóstico de tendencias en las bolsas de valores

- Industria: identificación de factores que afectan la calidad de sus productos

- Medicina: diagnóstico y tratamiento de enfermedades

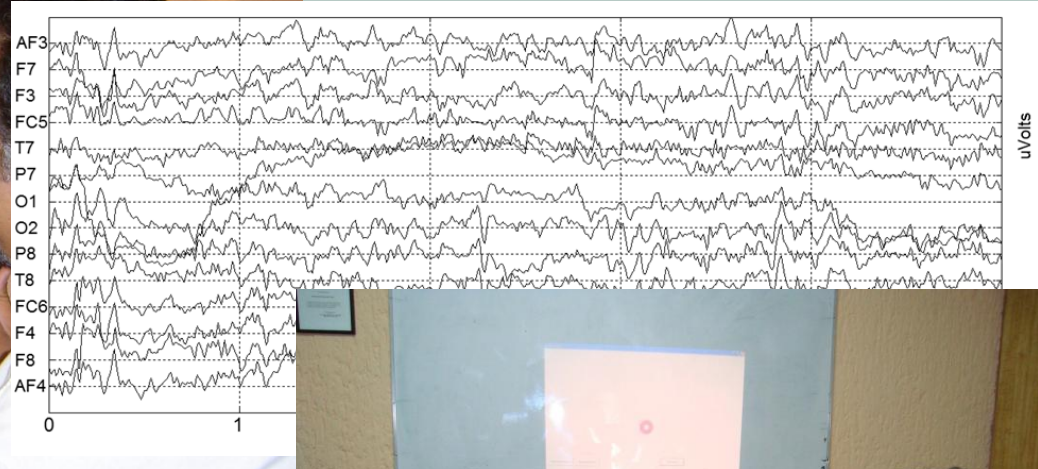
- etc...





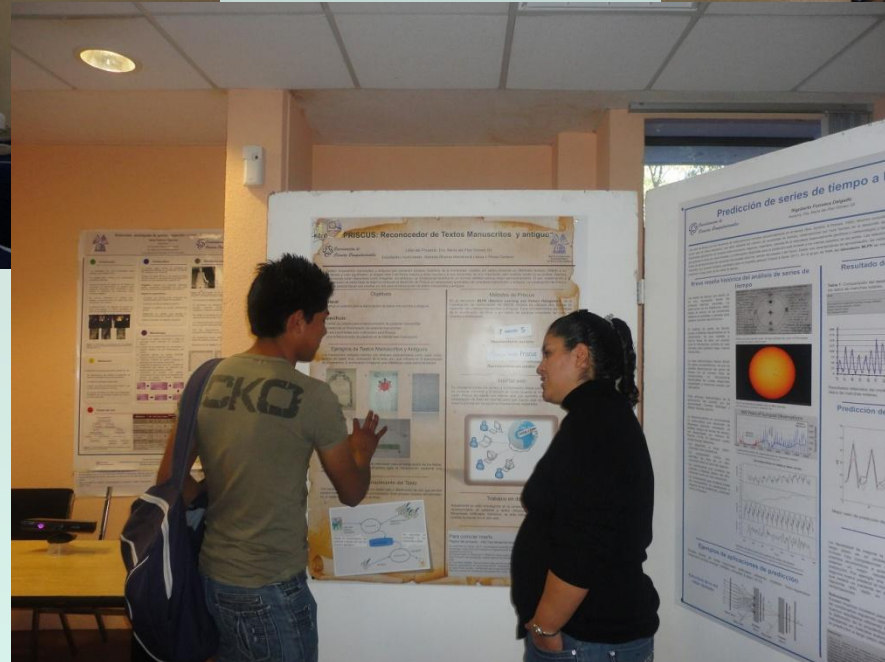
# Machine Learning and Pattern Recognition Lab

## Sensando Datos para experimentos...



Sensando datos para experimentos

# Integrantes del MLPR participando en seminarios...





# Dr. Francisco Martínez Trinidad, miembro del MLPR durante la ceremonia de entrega del Premio Estatal de Ciencia y Tecnología 2011

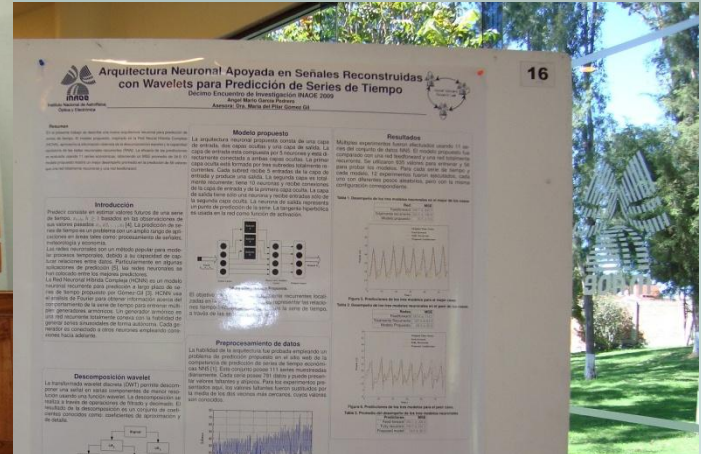


# Segundo lugar en el premio nacional de la bolsa mexicana de valores 2009, categoría de investigación, otorgado a la Dra. Pilar Gómez Gil y al Dr. Alfonso Mendoza, del Centro de Inteligencia de Negocios - UPAEP





# Los miembros del laboratorio participan en múltiples eventos con organizaciones, universidades y escuelas de la región, nacionales e internacionales



# La tecnología de información

- ◉ El concepto “TIC” en la actualidad cubre un gran número de facetas y diferentes áreas del conocimiento.
- ◉ México enfrenta problemas que pueden solucionarse a través de tecnologías de información en muy diversas áreas:
  - > Toma de decisiones
  - > Medicina
  - > Educación
  - > Seguridad
  - > Desarrollo urbano
  - > Desarrollo de tecnología propia

# Las TIC y sus herramientas de trabajo (1/2)

- La creación, uso y perfeccionamiento de tecnologías de información requiere de herramientas de muy diversos índoles:
  - > Abstractas (Matemáticas, Lógica)
  - > De comunicación (Teorías de redacción y estilos de expresión, difusión de conocimiento)
  - > De diseño (Modelos de Ingeniería de HW y Software, programación, redes)
  - > De gestión (Administración de Proyectos)
  - > De propósito específico (manejadores de bases de datos, lenguajes de programación, simuladores, optimizadores, clasificadores etc..)

# Las TIC's y sus herramientas de trabajo (2/2)

- Para resolver los retos que enfrentamos de manera no solo eficaz (que resuelven) sino también eficiente (que resuelven al mínimo costo) estas herramientas deben aplicarse de manera integral.
- Para usar las herramientas hay que conocerlas y perfeccionar su dominio de manera incremental.



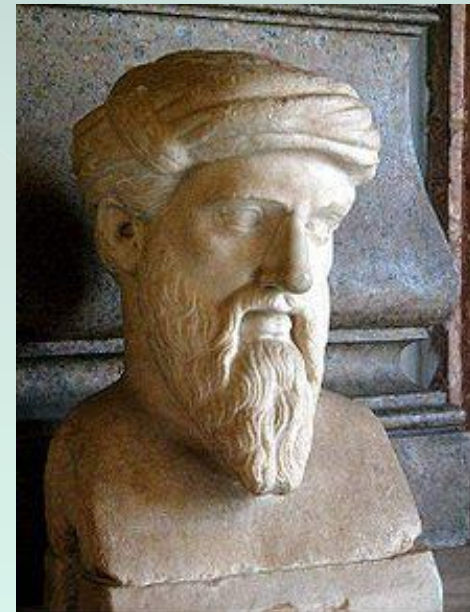
# Las TIC's y las Matemáticas

- Una de estas herramientas, las **matemáticas**, son de especial utilidad para la solución de un gran número de retos.
- Las matemáticas son herramientas que permiten construir otras de las herramientas nombradas anteriormente.

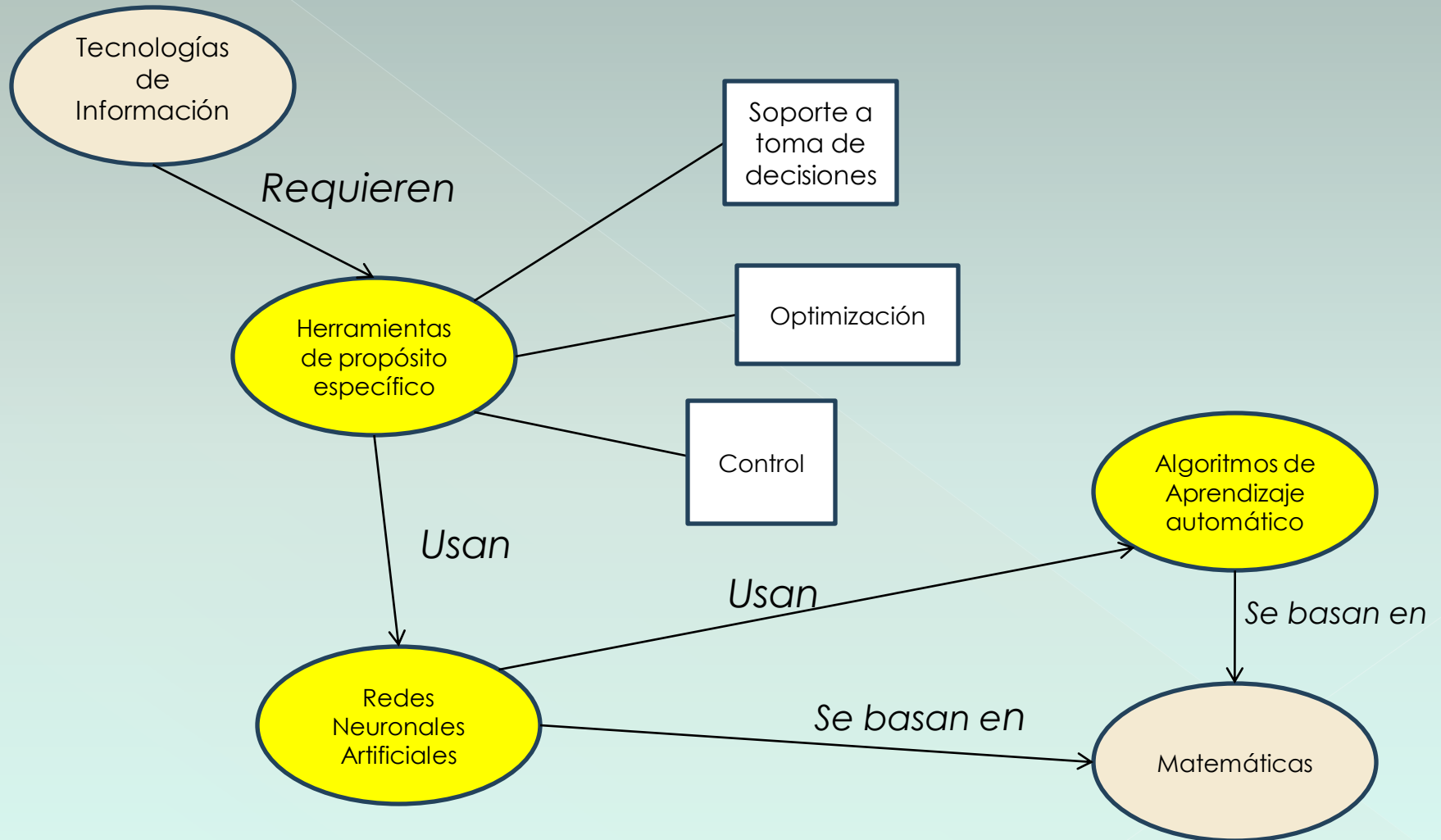
# Las matemáticas

- “Grupo de ciencias relacionadas que tratan con la lógica de la cantidad, la forma y el arreglo” (Wordweb Dictionary)

Pitágoras. Foto tomada de:  
<http://en.wikipedia.org/wiki/Mathematics>



# Una Relación (de tantas) entre TI y Matemáticas



# Las Redes Neuronales Artificiales (1/2)

- Son **modelos matemáticos** capaces de adaptar su comportamiento en respuesta a ejemplos presentados por el medio ambiente de manera supervisada o no supervisada (esto es **aprendizaje automático!**)

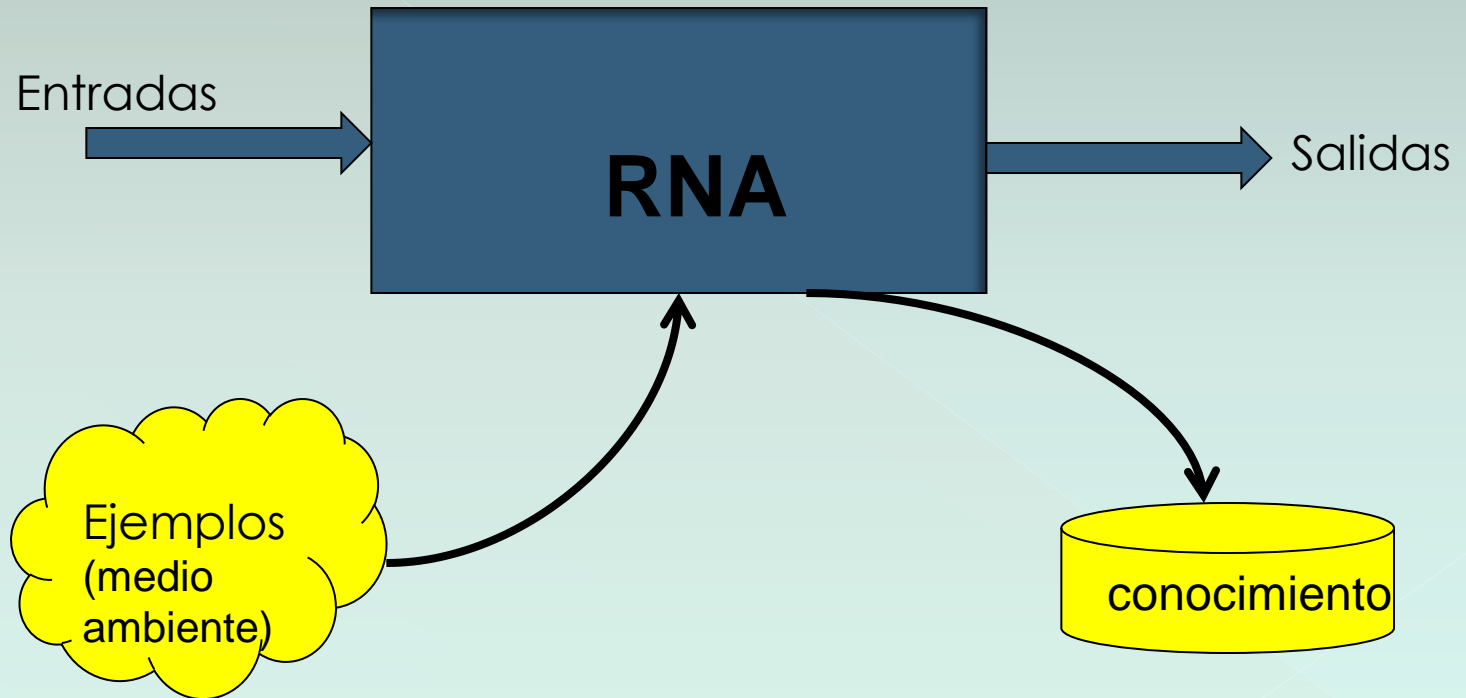


# Las Redes Neuronales Artificiales (1/2)

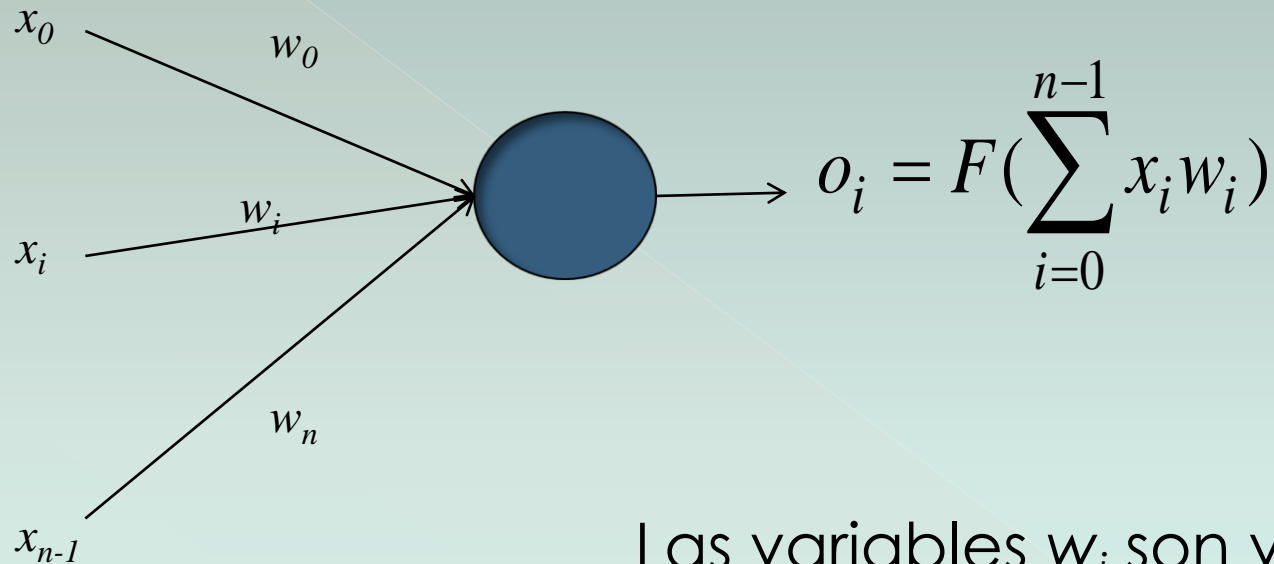
- Están inspiradas en la construcción del cerebro y las neuronas biológicas.
- Una red neuronal artificial es un procesador paralelo y distribuido, hecho de varios procesadores simples, que puede almacenar y utilizar conocimiento adquirido de la experiencia (Haykin 2009).

# El Contexto de Redes Neuronales

(Gómez-Gil, 2009)



# El componente fundamental: neurona



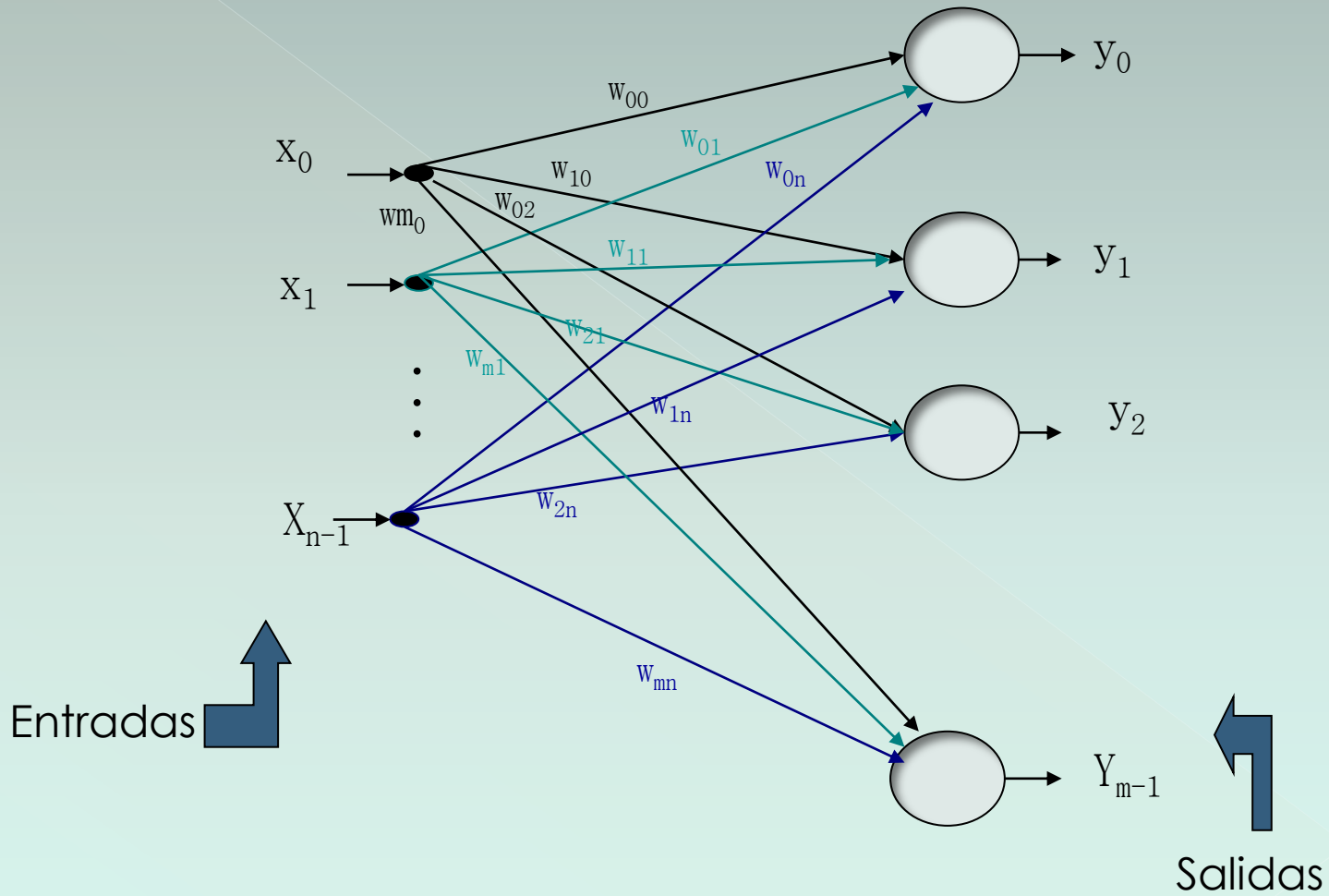
Las variables  $w_i$  son valores reales que contienen el conocimiento de la red neuronal

# Topologías principales de redes neuronales artificiales

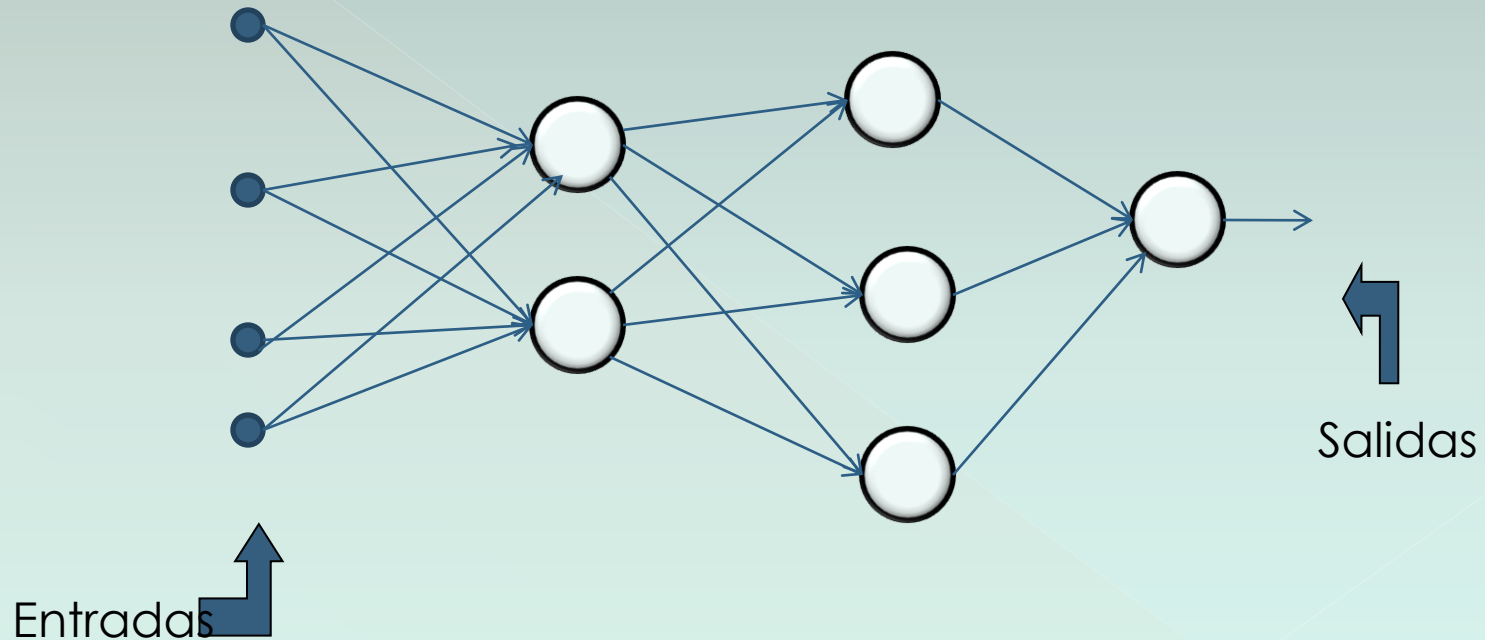
- El modelo paralelo y distribuido de una red neuronal puede tener varias formas, conocidas como:
  - > Redes de un nivel
  - > Redes de varios niveles
  - > Redes recurrentes



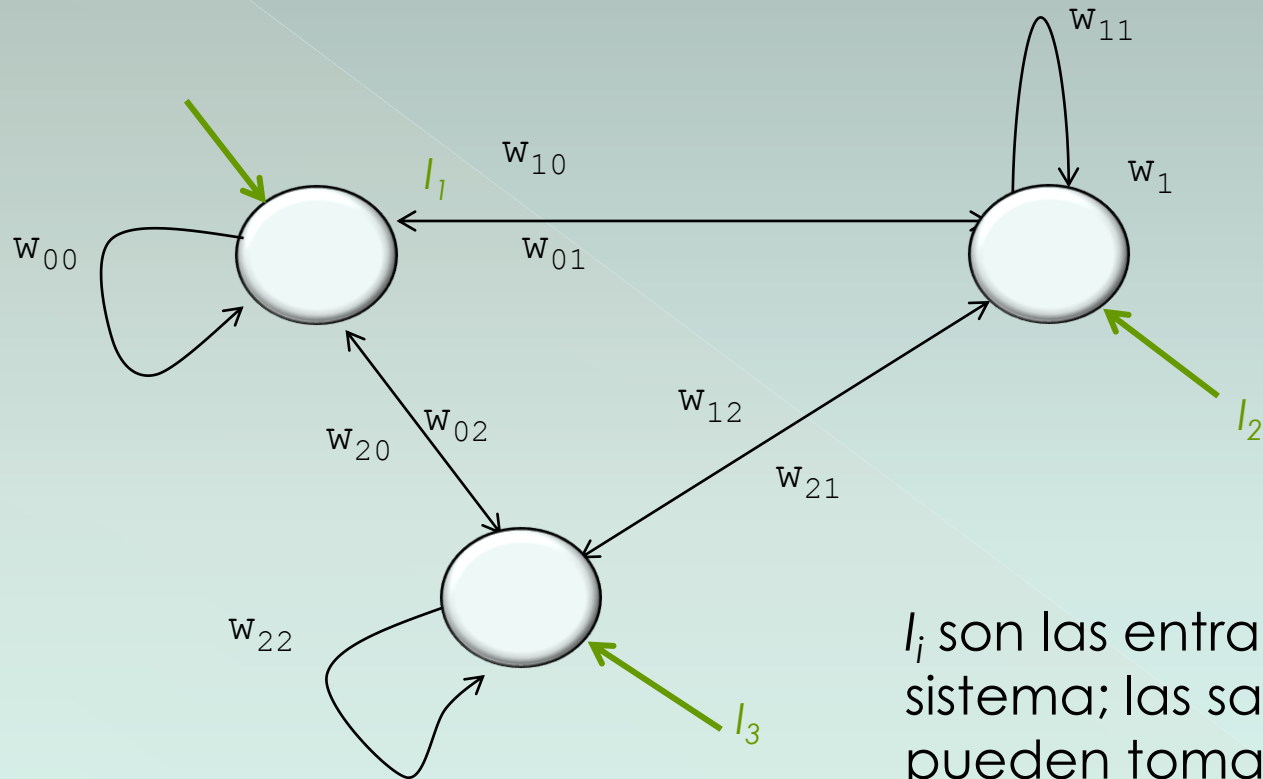
# Redes de un nivel



# Redes de varios niveles



# Redes Recurrentes



$I_i$  son las entradas al sistema; las salidas pueden tomarse de cualquier neurona

# Aprendizaje basado en ejemplos

- Las redes neuronales son capaces de modificar su comportamiento en respuesta al medio ambiente, el cual es “presentado a la red” a través de ejemplos significativos del problema.
- Para una Red Neuronal, aprender significa modificar los valores de los pesos (números reales), siguiendo una estrategia determinada conocida como “algoritmo de aprendizaje.”

# Tipos de algoritmos de aprendizaje para RNA

- Basados en corrección de errores
- Basados en memoria
- Basados en la regla de Hebb
- Aprendizaje competitivo
- Aprendizaje Boltzmann



# Un ejemplo de Aprendizaje basado en corrección de error: Retropropagación

- ⦿ Es un modelo de entrenamiento **supervisado**, esto es, la red neuronal conoce explícitamente qué es lo que debe aprender
- ⦿ El algoritmo consiste en hallar los pesos que minimizan el cuadrado del error promedio entre la salida real y la salida deseada en la red. Es una técnica de gradiente descendente.
- ⦿ La función de error total  $ET$  a minimizar, es la suma acumulada de los errores generados por cada ejemplo del conjunto de entrenamiento:

# Minimización de Error

$$ET = \sum_p E_p \quad \text{donde :}$$

$$E_p = \frac{1}{2} \sum_{j=1}^n (T_{pj} - O_{pj})^2$$

$T_{pj}$  = Valor de salida deseado para el j-ésimo neurón de la última capa de la red.

$O_{pj}$  = Valor de salida real en el j-ésimo neurón de la última capa de la red.

# Algoritmo de Retro-propagación (Gómez-Gil, 2009)

Sean:

$W_{ji}$  = Peso que conecta hacia el neurón  $j$  en la capa  $k$  desde el neurón  $i$  en la capa  $k-1$ .

$F$  = Función de activación (continua y diferenciable)

$O_{pj}$  =  $F(\text{NET}_{pj})$  Salida del  $j$ -ésimo neurón.

$\text{NET}_{pj}$  =  $\sum W_{ji}O_{pi}$

Donde  $O_{pi}$  corresponde a la entrada de la red ( $X_i$ ) si el neurón  $i$  está en la primera capa de la red.

$\Delta pW_{ji}$  = Incremento en el peso  $W_{ji}$  provocado por el  $p$ -ésimo patrón.

# Algoritmo de Retro- propagación (cont.)

- 1.- Inicialice todos los pesos y valores umbrales de la red con número reales pequeños generados al azar.
- 2.- Repita lo siguiente hasta que el error ET del conjunto de entrenamiento sea aceptablemente pequeño, o alguna condición predeterminada de "fin de entrenamiento" sea verdadera:
  - 2.1 Por cada patrón  $p$  en el conjunto de entrenamiento:
    - 2.1.1 Lea el vector de entrada  $\mathbf{X}_p$  y el vector de la salida deseada  $\mathbf{T}_p$ .
    - 2.1.2 Calcule la salida de la red.
    - 2.1.3 Calcule el error  $E_p$  generado por el patrón  $p$ .



# Algoritmo de Retro-propagación (cont.)

2.1.4 Ajuste todos los pesos de la red aplicando la siguiente regla ( regla delta generalizada) a cada uno de los pesos  $W_{ji}$

$$W_{ji}(t+1) = W_{ji}(t) + \Delta p W_{ji}$$

donde

$$\Delta p W_{ji} = \eta \delta_{pj} O_{pi}$$

$\eta$  = Coeficiente de aprendizaje ( $0 < \eta < 1$ )

$\delta$  Se calcula de la siguiente manera:

a) Si  $j$  corresponde a un neurón en la capa de salida de la red:

$$\delta_{pj} = (T_{pj} - O_{pj}) F' (NET_{pj})$$

# Algoritmo de Retro-propagación (cont.)

si la función de activación es la sigmoide, entonces:

$$F'(x) = F(x)(1 - F(x)), \text{ y}$$
$$\delta_{pj} = (T_{pj} - O_{pj}) O_{pj} (1 - O_{pj})$$

b) Si  $j$  no corresponde a un neurón de salida, es decir, está en una de las capas escondidas entonces:

$$\delta_{pj} = F'(NET_{pj}) \sum_k \delta_{pk} w_{kj}$$

donde la sumatoria acumula el error propagado hacia atrás.

2.1.5 Regrese al paso 2.1

3.- Regrese al paso 2.

# ?De donde salió todo eso?

- Un pedazo de la explicación de deducción del algoritmo de retropropagación , tomada del libro de texto de (Haykin 1999):

In a manner similar to the LMS algorithm, the back-propagation algorithm applies a correction  $\Delta w_{ji}(n)$  to the synaptic weight  $w_{ji}(n)$ , which is proportional to the partial derivative  $\partial \mathcal{E}(n)/\partial w_{ji}(n)$ . According to the *chain rule* of calculus, we may express this gradient as:

$$\frac{\partial \mathcal{E}(n)}{\partial w_{ji}(n)} = \frac{\partial \mathcal{E}(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)} \quad (4.6)$$

# Algunos Ejemplos de Aplicaciones de RNA

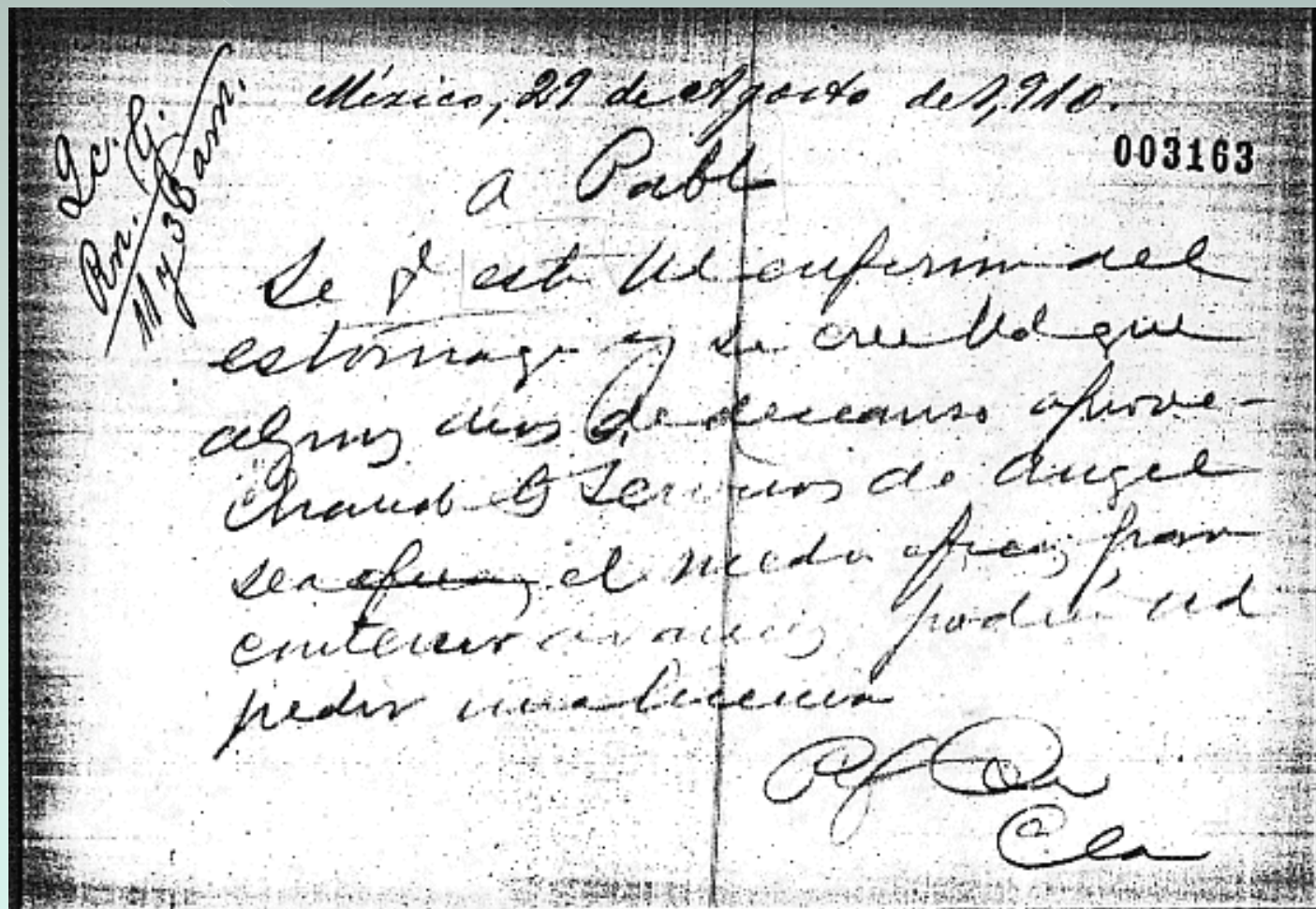
- Reconocimiento de caracteres manuscritos, impresos, de font antiguo, etc. .
- Construcción de Memorias asociativas.
- Reconocimiento de voz
- Control de robots
- Toma de decisiones administrativas, financieras etc.
- Reconocimiento de enfermedades
- Reconocimiento de señales de radio
- Predicción de Señales y series de tiempo Caóticas
- Generación de reglas para sistemas expertos
- Aplicaciones en economía para predicción
- Aplicaciones en geología, meteorología, astronomía
- Aplicaciones en ....

# Algunos de los problemas que nosotros estudiamos...

- Predicción de series de tiempo caóticas y de otros tipos utilizando Redes Neuronales Complejas
- **PRISCUS**: Reconocimiento automático de **escritura manuscrita** en documentos antiguos
- **Aproximadores** de funciones basados en redes neuronales para la estimación futura de
- **calificaciones crediticias** en gobiernos locales mexicanos
- Sistemas de **biometría multimodal** utilizando modelos adaptivos

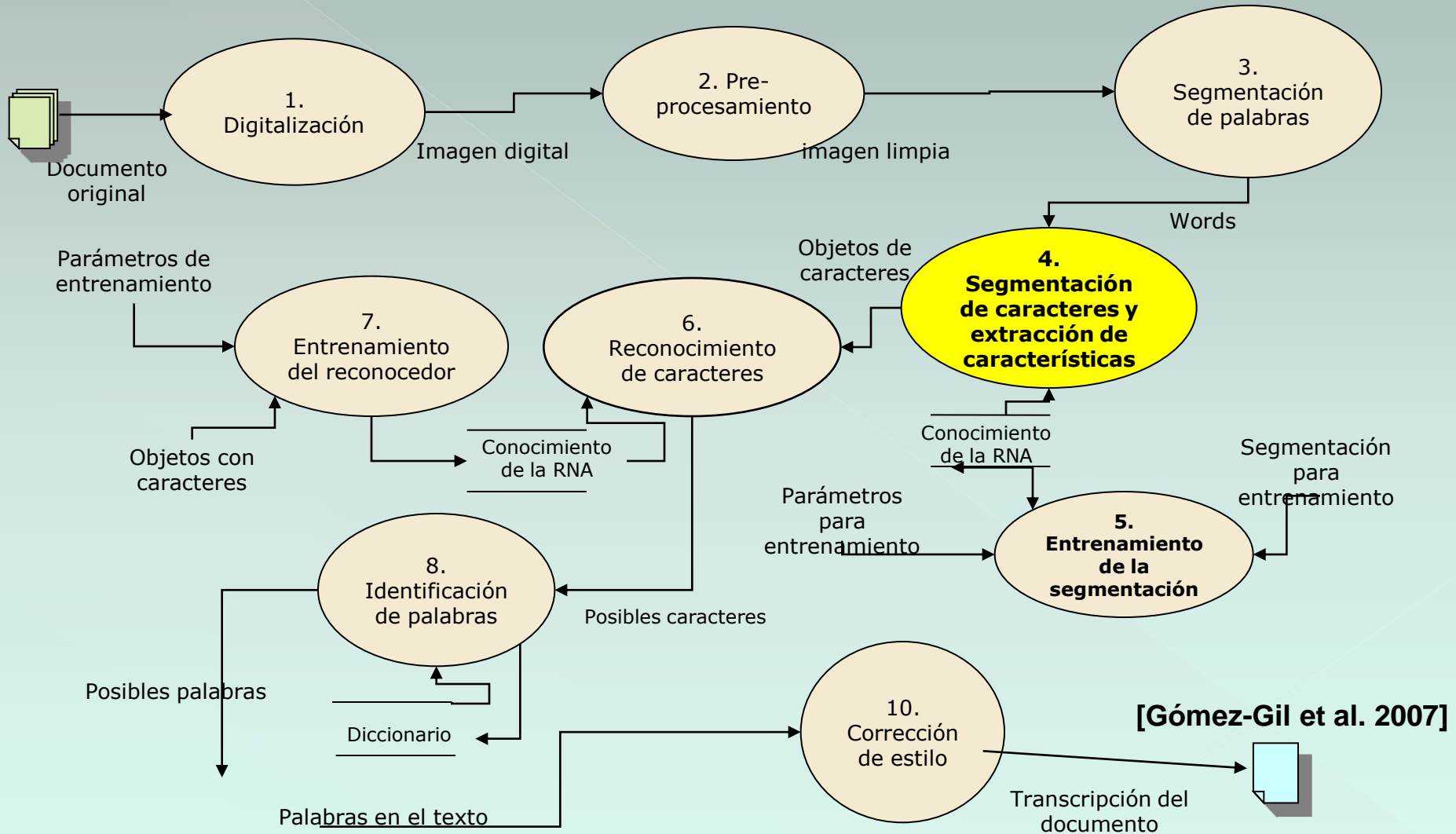


# El reconocimiento de manuscritos es un reto!



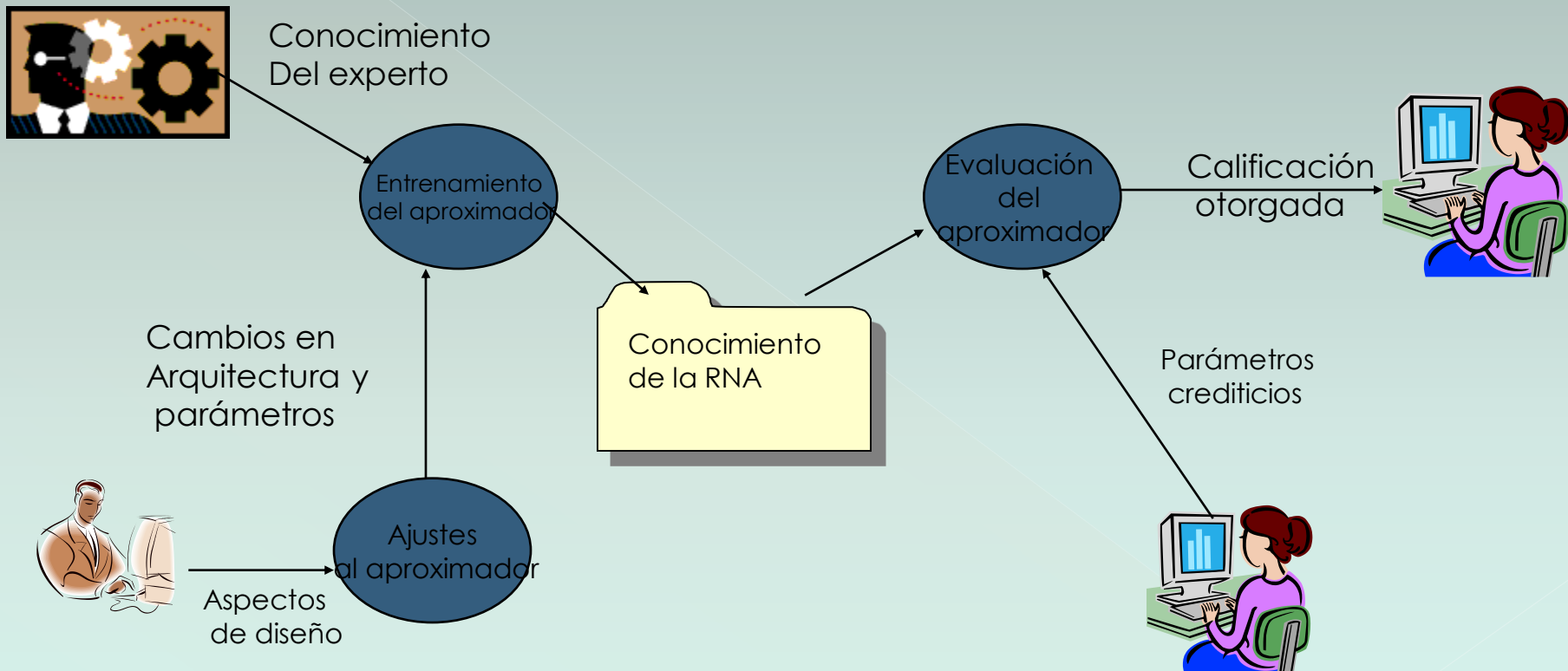
[Spinola & Linares 2000]

# Un OCR para documentos manuscritos

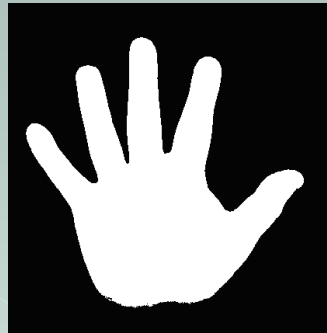


# Estructura de un aproximador crediticio

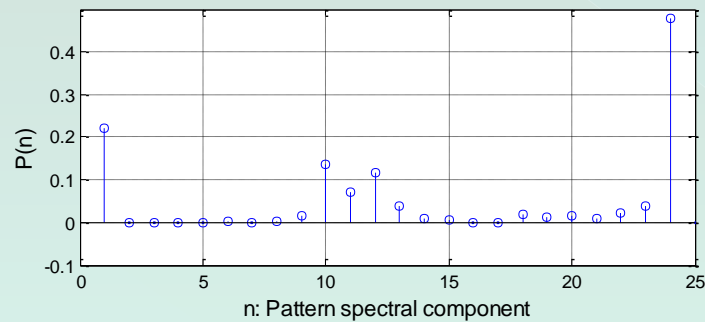
(Mendoza & Gómez-Gil 2010)



# Biometría multimodal (Ramírez-Cortés et al. 2008)

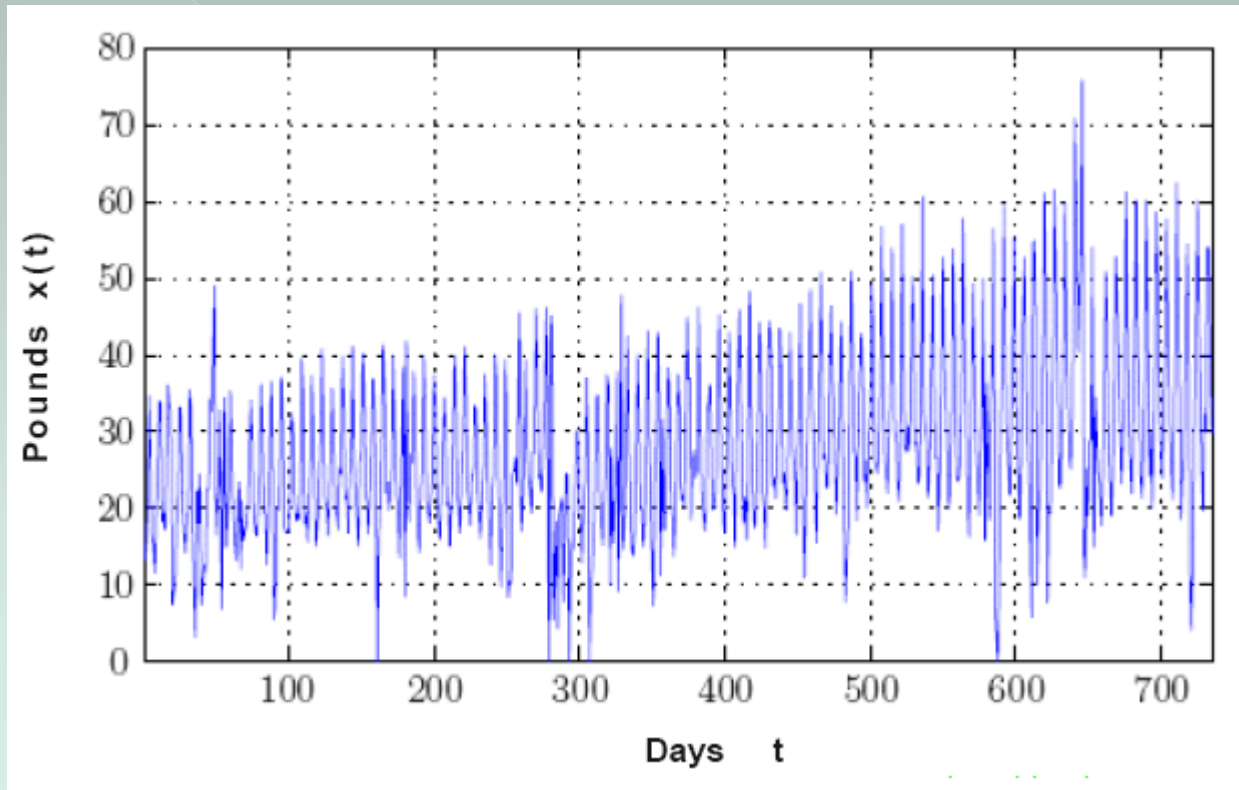


Extracción de características: Imágenes en niveles de gris e imagen binarizada



Extracción de características: Modelo pecstrum

# Predicción a largo plazo de señales caóticas (García-Pedrero & Gómez-Gil 2010)



Serie de tiempo NN5-1 de la NN5 Forecasting Competition for Artificial Neural Networks and Computational Intelligence. (Crone, 2008).



Para conocer un poco mas de estos y otros proyectos consulta:

<http://ccc.inaoep.mx/~pgomez/publicaciones.htm>

# Conclusiones

- Las TI son una excelente oportunidad para solucionar una gran cantidad de problemas actuales
- Su uso eficiente y eficaz requiere de conocer y dominar varias herramientas, entre ellas las matemáticas
- Las matemáticas son una herramienta para la construcción de modelos basados en aprendizaje automático y redes neuronales artificiales
- Las Redes Neuronales Artificiales tienen aplicaciones de diversas áreas donde TI se aplica continuamente

# Muchas gracias por su atención!

Estas diapositivas están disponibles en:

<http://ccc.inaoep.mx/~pgomez/VeranoINAOE-13.pdf>

[pgomez@acm.org](mailto:pgomez@acm.org)

[pgomez@inaoep.mx](mailto:pgomez@inaoep.mx)

# Bibliografía

- Crone S.F.: "NN5 Forecasting Competition for Artificial Neural Networks & Computational Intelligence." Available since 2008 at: <http://www.neural-corecasting-competition.com/>. Last consulted at March, 2009. (2008).
- García-Pedrero, A and P. Gómez-Gil. "Time Series Forecasting using Recurrent Neural Networks and Wavelet Reconstructed Signals". Proceedings of the 20th. International Conference on Electronics, Communications and Computers. CONIELECOMP 2010.
- Gómez-Gil, P. Tutorial: [An introduction to the use of Artificial Neural Networks](#). 8<sup>th</sup>. Mexican International Conference on Artificial Intelligence. MICAI 2009. Disponible en: [http://ccc.inaoep.mx/~pgomez/cursos/redes%20neuronales%20artificiales/index\\_archivos/Page530.htm](http://ccc.inaoep.mx/~pgomez/cursos/redes%20neuronales%20artificiales/index_archivos/Page530.htm)
- Gómez-Gil, P. De-Los-Santos Torres G., Navarrete-García J. Ramírez-Cortés M.. "[The Role of Neural Networks in the interpretation of Antique Handwritten Documents .](#)" in *Hybrid Intelligent Systems. Analysis and Design Series: Studies at Fuzziness and Soft Computing*. Vol . 208. Editors: Castillo, O. Melin, P. Kacprzyk W. Springer. ISBN-10: 3-540-37419-1. Pags. 269-281. 2007
- Mendoza, A. y Gómez-Gil, P. "Herramientas para el Pronóstico de la Calificación Crediticia en las Finanzas Públicas Estatales en México: Redes Neuronales Artificiales, Modelo Probit Ordenado y Análisis Discriminante." A publicarse en *Herramientas de Diagnóstico y Respuesta de las Finanzas Públicas Locales en México en un Entorno de Crisis*" Editor: Mendoza, A. Editorial Porrúa. 2010.
- S. Haykin. *Neural Networks, A comprehensive Foundation*. MacMillan College Publishing Company. 1999
- Ramírez-Cortés, JM., P. Gómez-Gil, Sánchez-Perez, G. Prieto-Castro, C. "[A shape-based hand recognition approach using the pattern spectrum.](#)" *Journal of Electronic Imaging*. Jan-March 2009. Vol. 18, No. 1. ISSN 1017-9909. Digital version: Vol. 18, 013012 (2009); DOI:10.1117/1.3099712