



**Universidad Politécnica de Madrid**

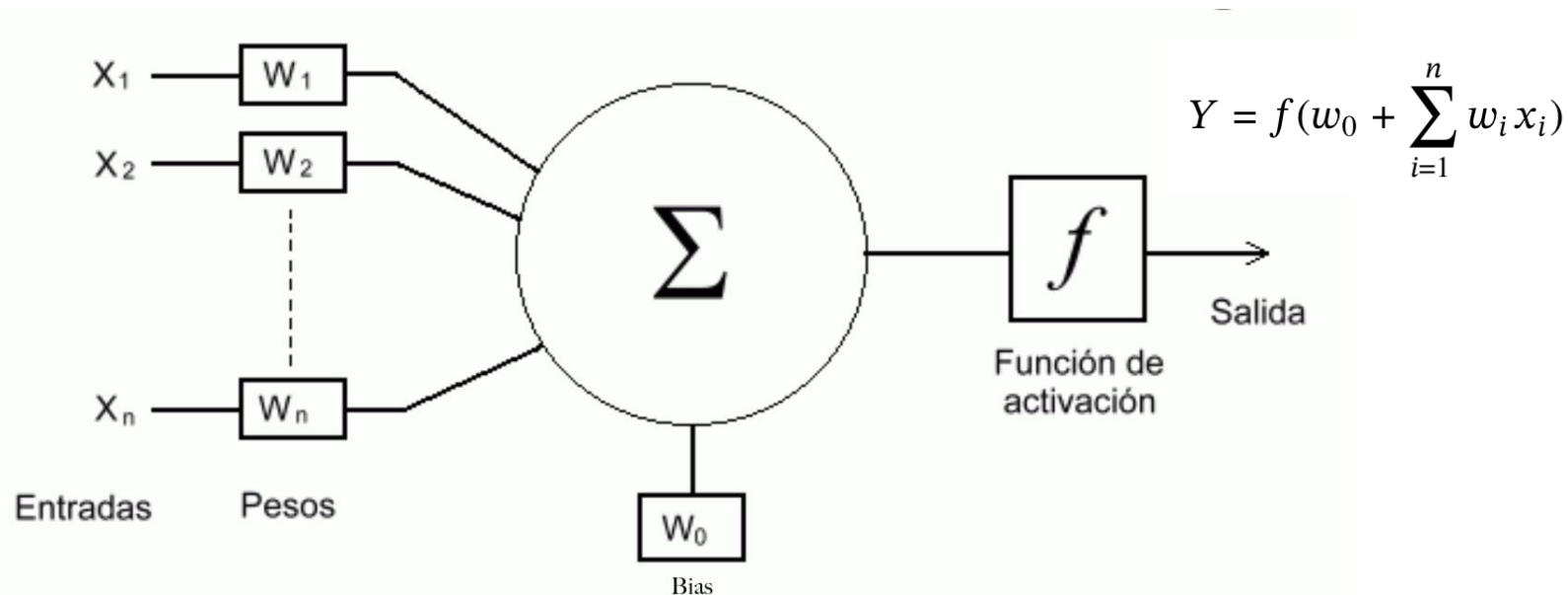
**Introducción a la inteligencia artificial generativa  
a través de los grandes modelos de lenguaje**

***Introducción IA (Parte II)***

**Javier Conde  
Pedro Reviriego**

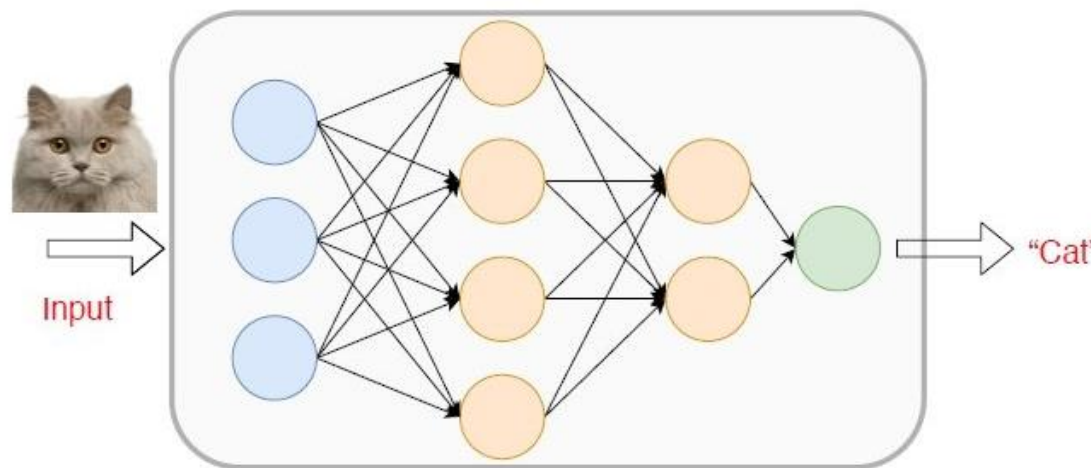
# Aprendizaje supervisado: redes neuronales y Deep Learning

- Dentro del aprendizaje supervisado una de las técnicas que ofrece mejores resultados es el aprendizaje profundo basado en redes neuronales conocido como Deep Learning
- Una neurona recibe múltiples datos de entrada y genera una salida
- La función de activación proporciona no linealidad



# Aprendizaje supervisado: redes neuronales y Deep Learning

- Las neuronas se pueden combinar formando redes neuronales complejas con múltiples entradas, múltiples capas y múltiples salidas. Se pueden procesar imágenes, texto, vídeo, etc.
- Deep Learning se refiere a redes NN con muchas capas
- El entrenamiento se basa en ajustar los pesos de las neuronas
- La base del entrenamiento es el algoritmo de *backpropagation*
- Diferentes arquitecturas de redes neuronales para diferentes casos de uso



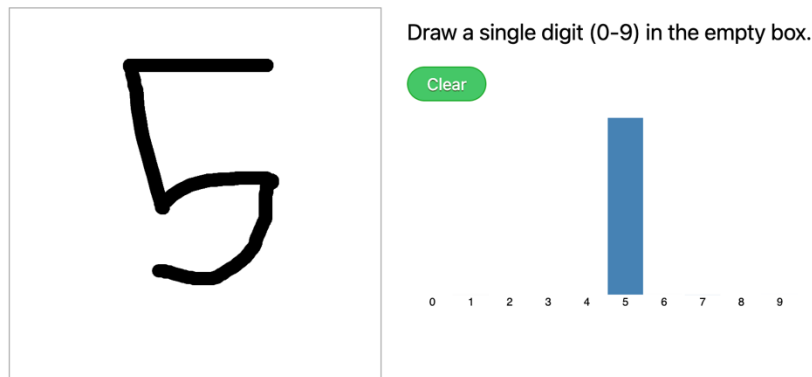
# Aprendizaje supervisado: redes neuronales y Deep Learning

- Ejemplo: red neuronal capaz de reconocer dígitos manuscritos:

<https://mco-mnist-draw-rwpkka3zaa-ue.a.run.app/>

 TensorFlow.js + *MNIST*

Draw a digit by hand and have it classified in real time, as you draw, by a machine learning model trained against the [MNIST](#) dataset. The model used for this page is a convolutional neural network (CNN) built using Keras/TensorFlow on a Google Tensor Processing Unit (TPU). It's explained in depth and is available for you to build for yourself at [bit.ly/mco-mnist-lab](https://bit.ly/mco-mnist-lab). Source code on [Github](#).

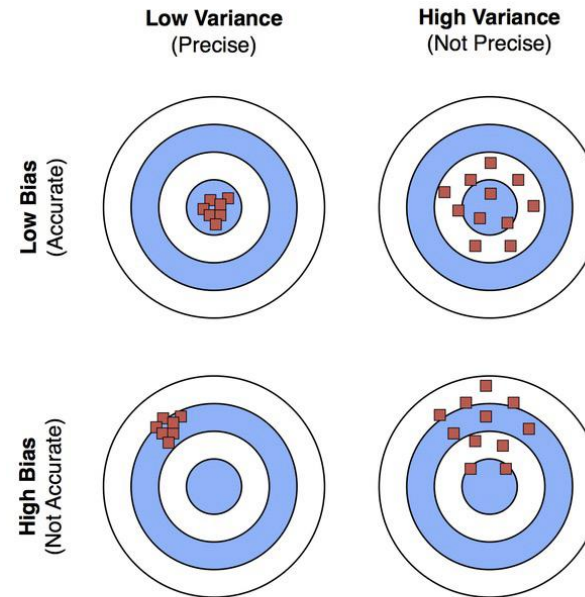



Made with ❤ by [Marc](#).

- Un aspecto fundamental de la IA es la evaluación de los modelos una vez entrenados
- Típicamente 80% datos para entrenamiento y 20% para evaluación de los modelos
- Diferentes métricas de evaluación
- Dependiendo del problema nos interesará maximizar unas métricas de evaluación u otros
- Los datos de evaluación no pueden usarse en el entrenamiento para evitar overfitting

## Diferentes métricas para evaluar la IA

- Ejemplo regresión: error cuadrático medio (MSE)
- Ejemplo clasificación: Matriz de confusión
- **TP** (True Positives / Verdaderos positivos): Casos correctamente clasificados como positivos.
- **TN** (True Negatives / Verdaderos negativos): Casos correctamente clasificados como negativos
- **FP** (False Positives / Falsos positivos): Casos incorrectamente clasificados como positivos (falsas alarmas)
- **FN** (False Negatives / Falsos negativos): Casos incorrectamente clasificados como negativos (fallos en detección)
- Métricas: accuracy, precisión, recall



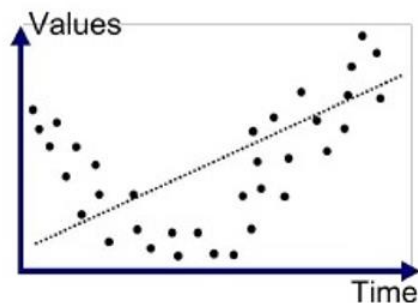
 This work by Sebastian Raschka is licensed under a Creative Commons Attribution 4.0 International License.

		Real Label		
		Positive	Negative	
Predicted Label	Positive	True Positive (TP)	False Positive (FP)	Precision = $\frac{\sum TP}{\sum TP + FP}$
	Negative	False Negative (FN)	True Negative (TN)	
		Recall = $\frac{\sum TP}{\sum TP + FN}$		Accuracy = $\frac{\sum TP + TN}{\sum TP + FP + FN + TN}$

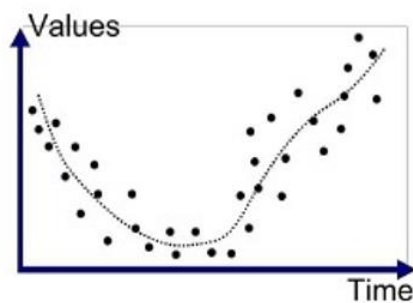


# Underfitting vs overfitting

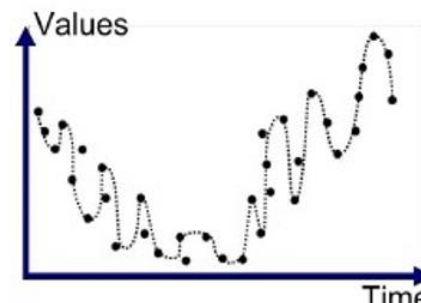
- Lo ideal es que el modelo entrenado sea capaz de generalizar la solución del problema con los datos del problema
- Overfitting (sobreajuste): el modelo se adapta demasiado a los datos de entrenamiento, por lo que disminuye su rendimiento con los datos de evaluación
- Underfitting (subajuste): el modelo no tiene suficientes datos por lo que no es capaz de generalizar el problema por completo



Underfitted



Good Fit/Robust



Overfitted

- **Aspectos técnicos:** diseño de modelos, su rendimiento, y los sesgos que pueden surgir en los sistemas de IA.
- **Aspectos legales:** seguridad, privacidad, la posibilidad de suplantaciones, los riesgos asociados a tecnologías como los deepfakes.
- **Impactos:** impacto de decisiones tomadas por sistemas de IA, los dilemas éticos que plantean, la necesidad de garantizar transparencia y falta de explicabilidad en su funcionamiento.
- **Impactos sociales:** desplazamiento de empleos, prejuicios, discriminación, riesgos para la seguridad física, pérdida de la conexión humana, y el uso de la IA para desinformar o manipular.
- **Reflexiones filosóficas:** dependencia excesiva de la IA, el empobrecimiento intelectual, la pérdida de creatividad o la deshumanización que podría generarse con el uso extendido de estas tecnologías.
- **Regulación:** equilibrio en su regulación y el aprovechamiento de sus capacidades.





**Universidad Politécnica de Madrid**

**Introducción a la inteligencia artificial generativa  
a través de los grandes modelos de lenguaje**

***Introducción IA (Parte II)***

**Javier Conde  
Pedro Reviriego**