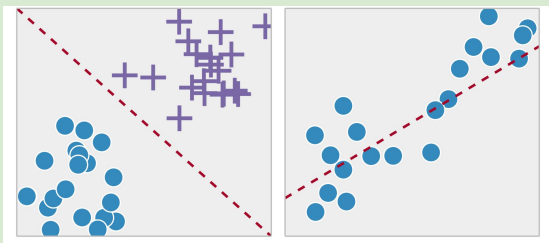


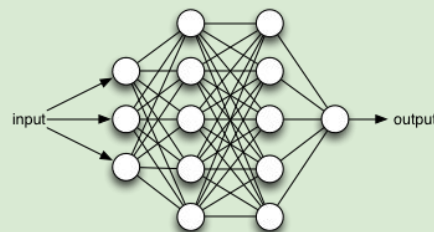
# Redes Neuronales Profundas

Lucas C. Uzal  
Guillermo L. Grinblat

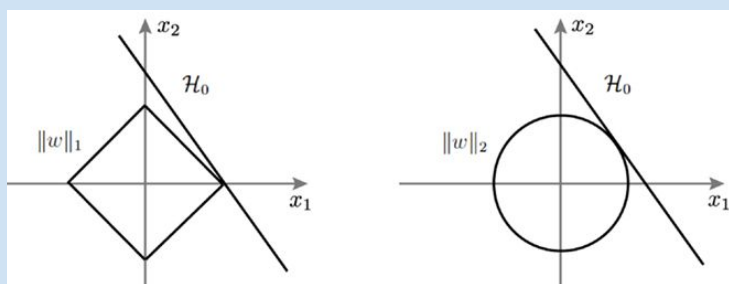
## Unidad 1: Intro Machine Learning



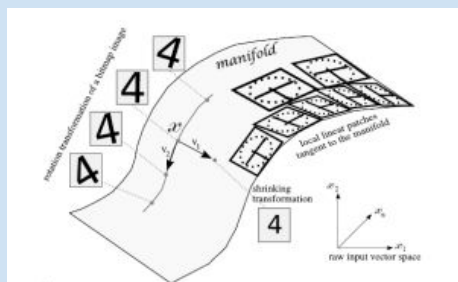
## Unidad 2: Artificial Neural Networks



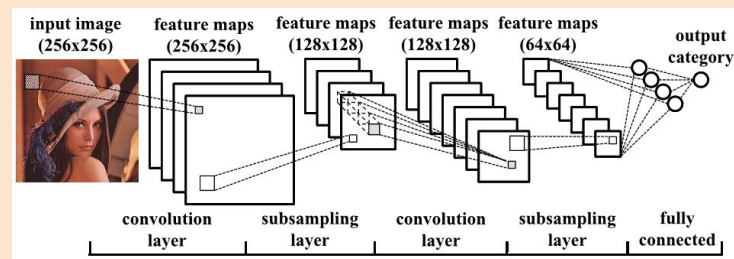
## Unidad 3: Técnicas de Regularización



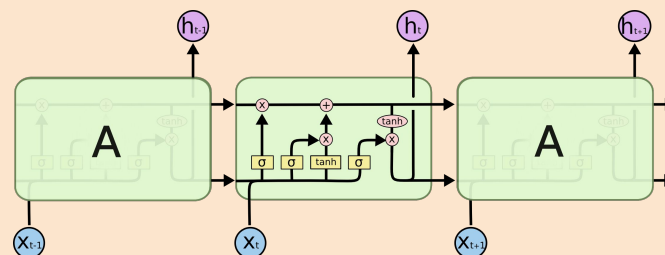
## Unidad 5: Aprendizaje de Representaciones



## Unidad 4: Convolutional Neural Networks



## Unidad 6: Recurrent Neural Networks



## Herramientas y Aplicaciones:



theano

python™

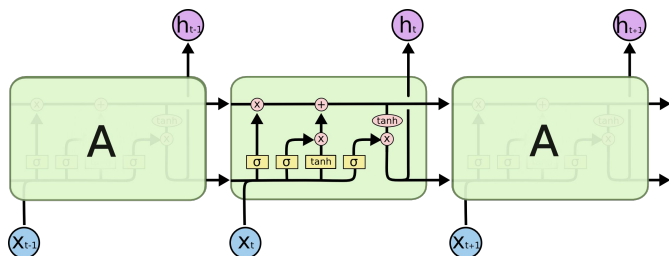
# Práctica

## Wikicorpus Español

TP4

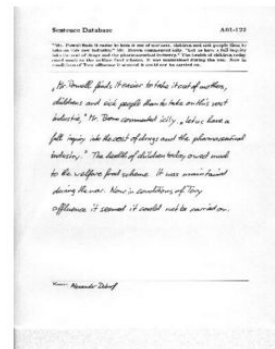


WIKIPEDIA  
La enciclopedia libre

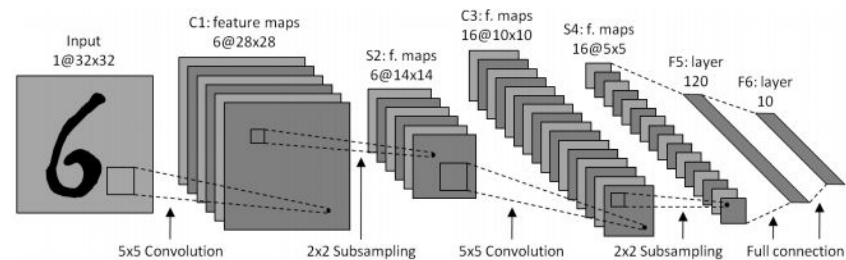


## IAM Handwriting Database

TP3



industrie, Mr. Bonn commented icily, "let us have a  
industrie icily  
Let have



uno de los puntos del espacio que contienen todos

# Bibliografía

## Deep Learning

An MIT Press book

Ian Goodfellow, Yoshua Bengio and Aaron Courville

<http://www.deeplearningbook.org/>

Hyvärinen, A., Hurri, J., & Hoyer, P. O. (2009). *Natural Image Statistics: A Probabilistic Approach to Early Computational Vision* (Vol. 39). Springer Science & Business Media.

Christopher M. Bishop (2006) *Pattern Recognition and Machine Learning*. Springer.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.

J. Schmidhuber. Deep Learning in Neural Networks: An Overview. *Neural Networks*, Volume 61, January 2015, Pages 85-117 (DOI: 10.1016/j.neunet.2014.09.003)

Theano Documentation. <http://deeplearning.net/software/theano/>

TensorFlow Documentation. <https://www.tensorflow.org/>

Keras Documentation. <http://keras.io/>

Caffe Documentation. <http://caffe.berkeleyvision.org/>

# Unidad 1: Introducción al Aprendizaje Automatizado

Curso: Redes Neuronales Profundas

# ¿Qué es Machine Learning?

*“A computer program is said to learn from experience  $E$  with respect to some class of tasks  $T$  and performance measure  $P$ , if its performance at tasks in  $T$ , as measured by  $P$ , improves with experience  $E$ .” [Mitchell 1997]*



Machine Learning  
Mitchell, T.M.  
1997  
McGraw-Hill

“Se dice que  
un programa de computadora aprende  
de la experiencia  $E$   
sobre un tipo de tareas  $T$   
y medida de desempeño  $P$ ,  
si su desempeño en la tareas de tipo  $T$   
medido en términos de  $P$ ,  
mejora con la experiencia  $E$ ”  
[Mitchell 1997]

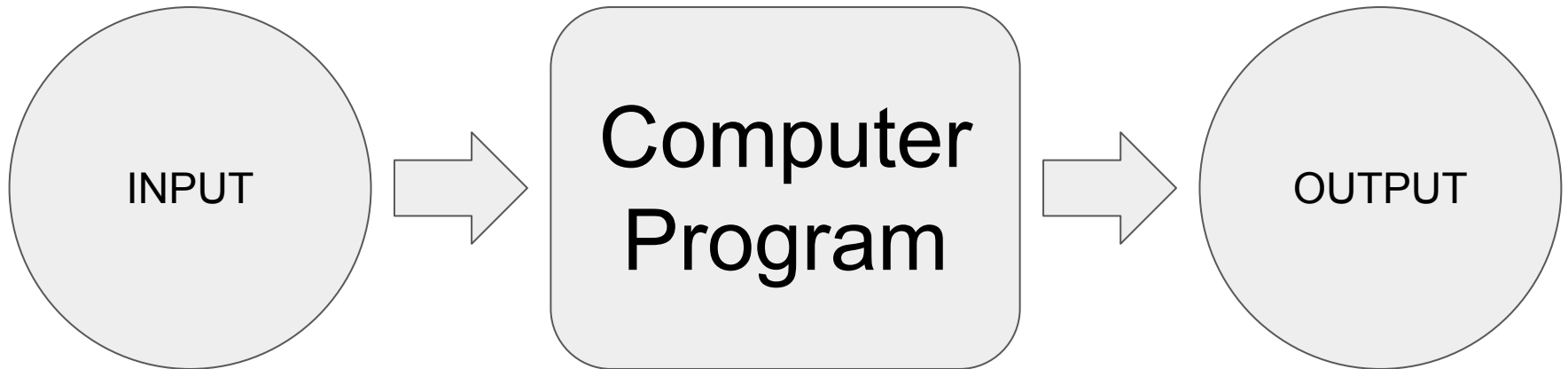
**EXPERIENCE ✓**

**TASK ✓**

**PERFORMANCE**

**MEASURE ✓**

# Machine Learning Tasks



# Machine Learning Tasks

Clasificación

Estimación de densidades

Detección de anomalías

Regresión

Transcripción

Traducción

Filtrar Ruido

Predicción de outputs  
estructurados

Clasificación con entradas  
faltantes

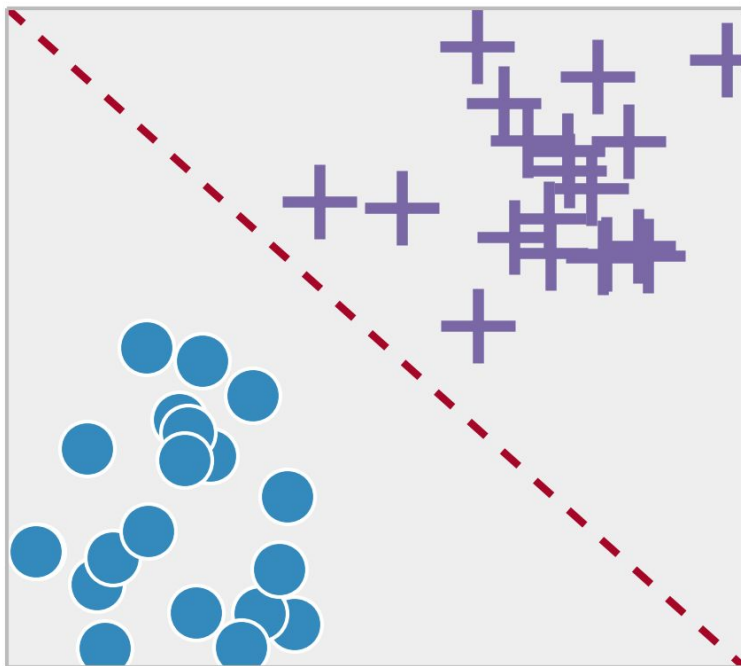
Recuperar información  
perdida

Generar datos sintéticos

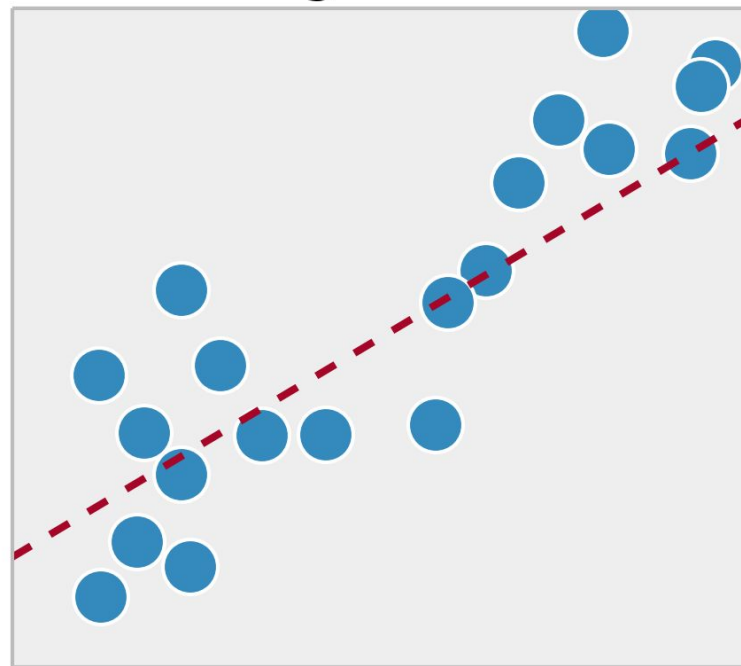


# Clasificación y Regresión

Classification



Regression





**mite**

**container ship**

**motor scooter**

**leopard**

	mite		container ship		motor scooter		leopard
	black widow		lifeboat		go-kart		jaguar
	cockroach		amphibian		moped		cheetah
	tick		fireboat		bumper car		snow leopard
	starfish		drilling platform		golfcart		Egyptian cat



**grille**

**mushroom**

**cherry**

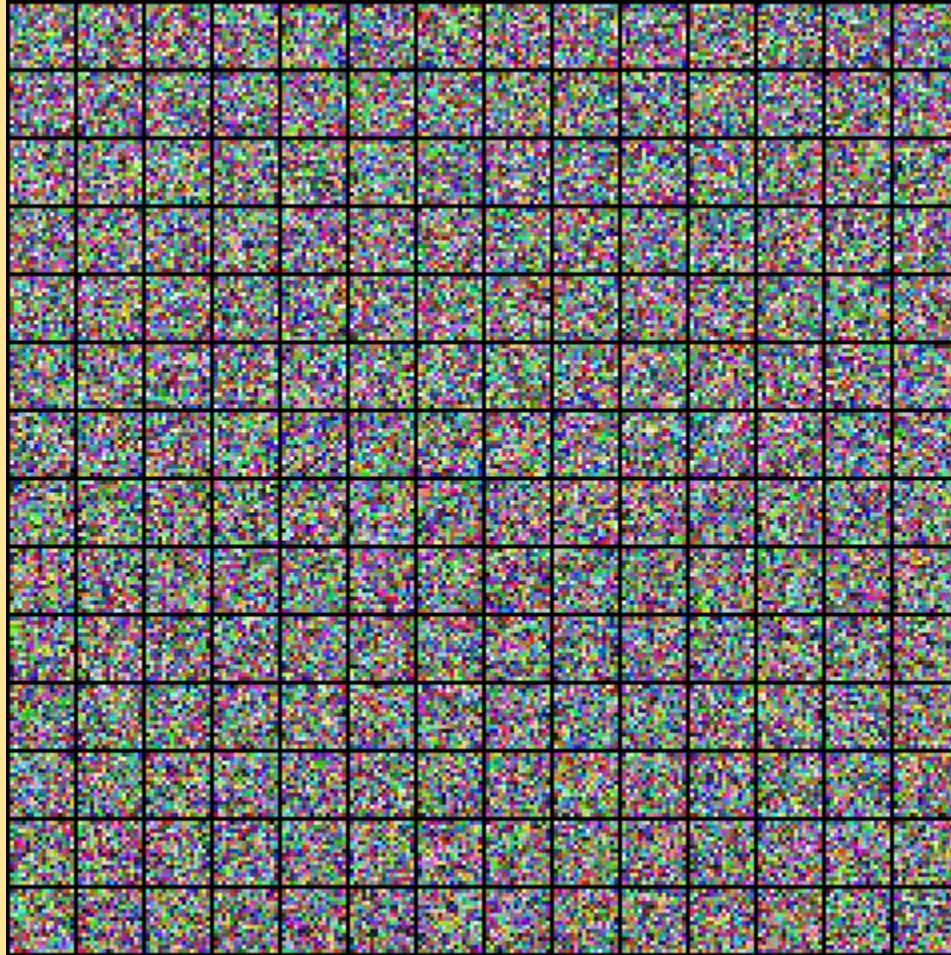
**Madagascar cat**

	convertible		agaric		dalmatian		squirrel monkey
	grille		mushroom		grape		spider monkey
	pickup		jelly fungus		elderberry		titi
	beach wagon		gill fungus		ffordshire bullterrier		indri
	fire engine		dead-man's-fingers		currant		howler monkey

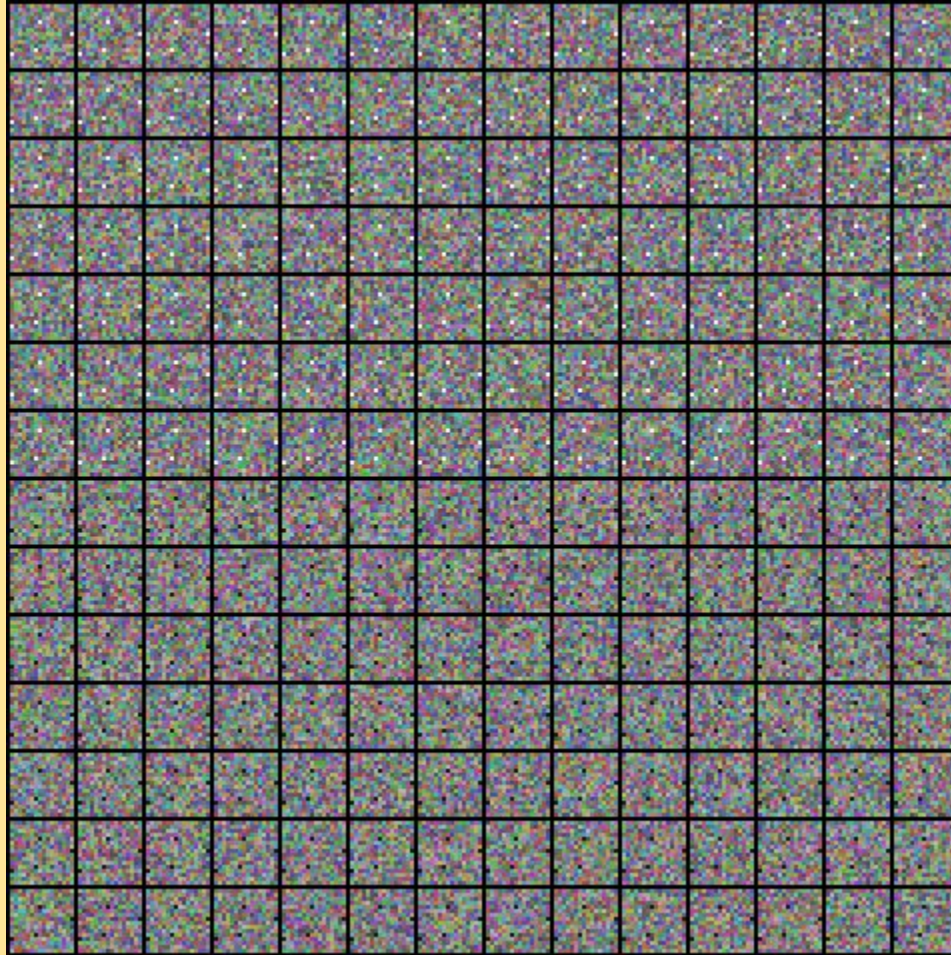
Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.



**¿Qué se distingue acá?**

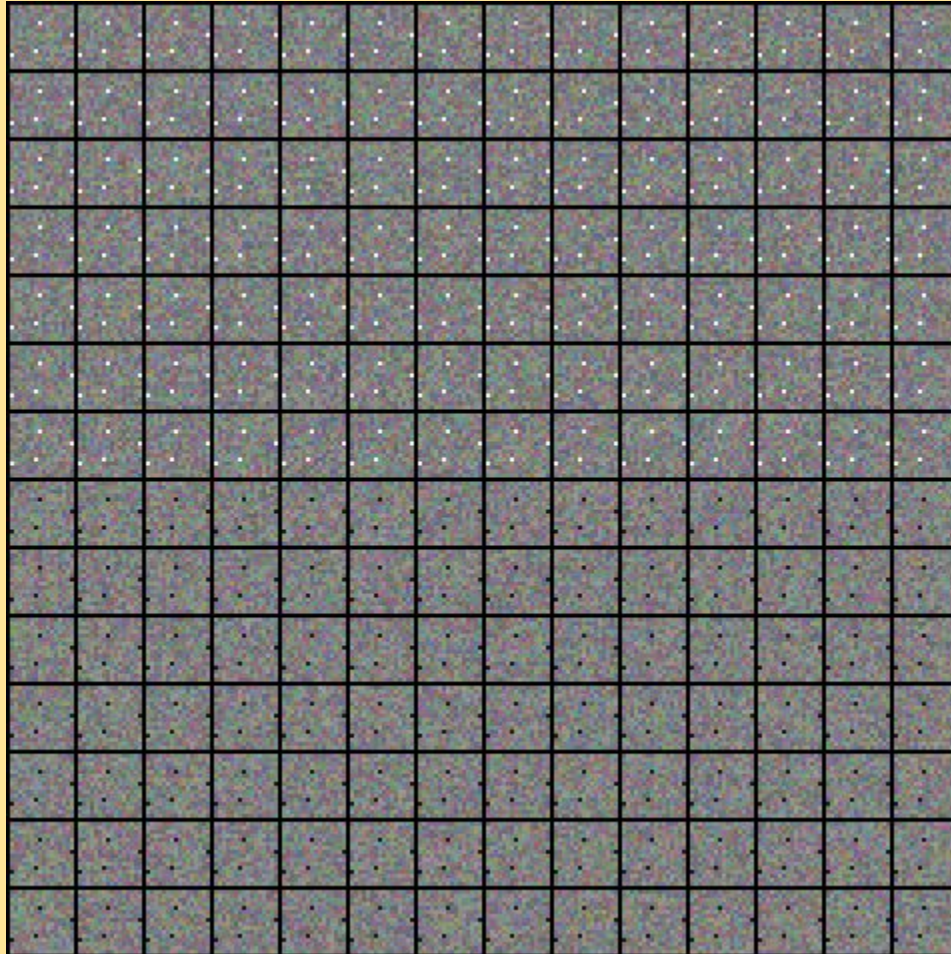


**¿Qué se distingue acá?**





**¿Qué se distingue acá?**



# **Cuándo pensar en Deep Learning**

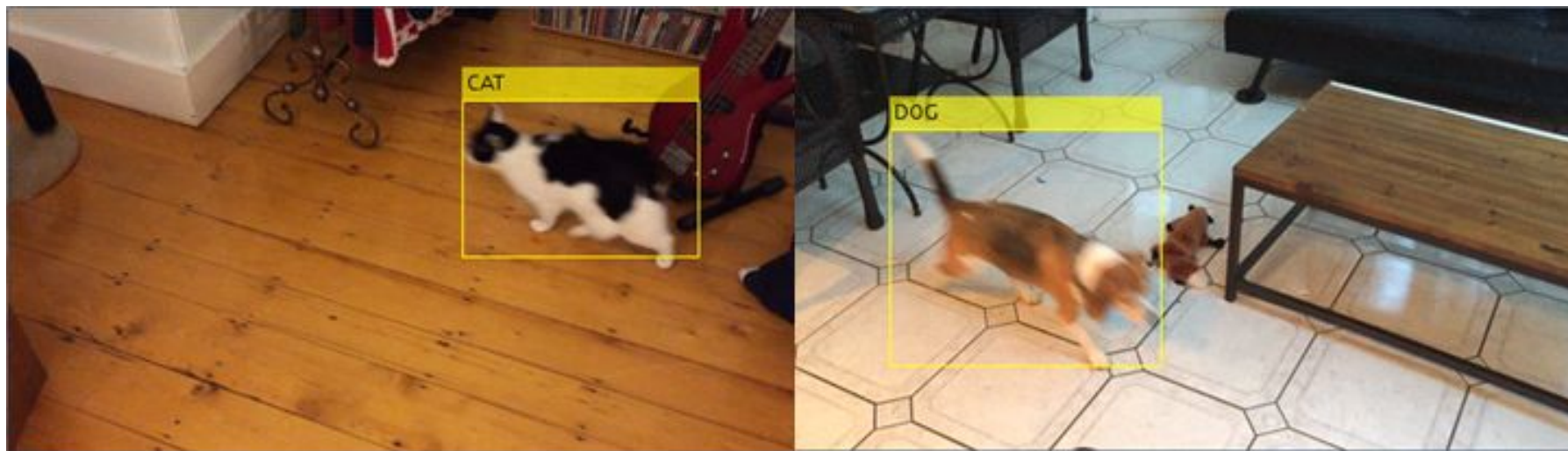
Muchas entradas

Complejidad

Variables de alto  
nivel de abstracción

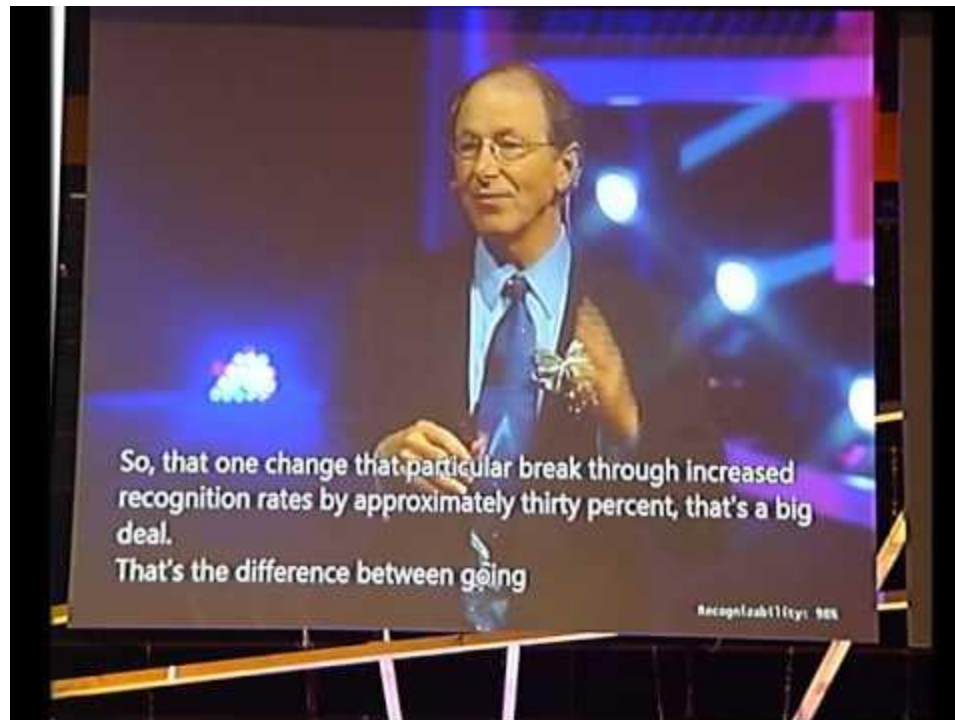
Muchos datos

# Regresión: coordenadas del recuadro



*bounding box*

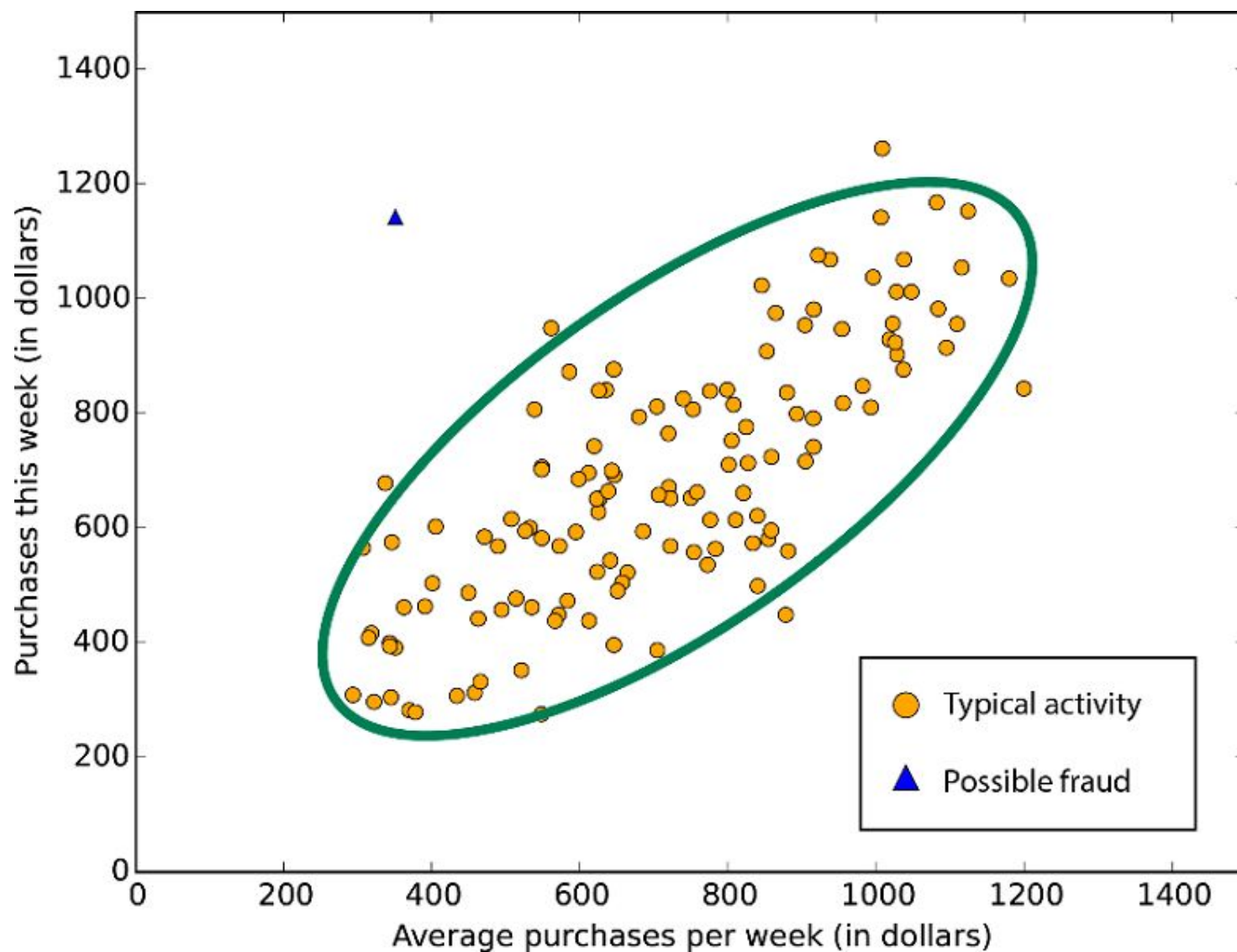
# Transcripción y traducción simultáneas



Microsoft 2012



# Detección de anomalías



# Generación de datos sintéticos. GAN



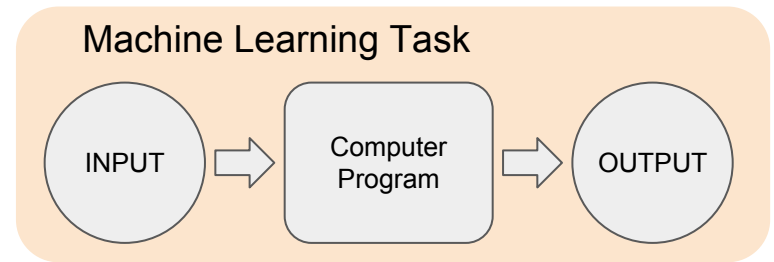
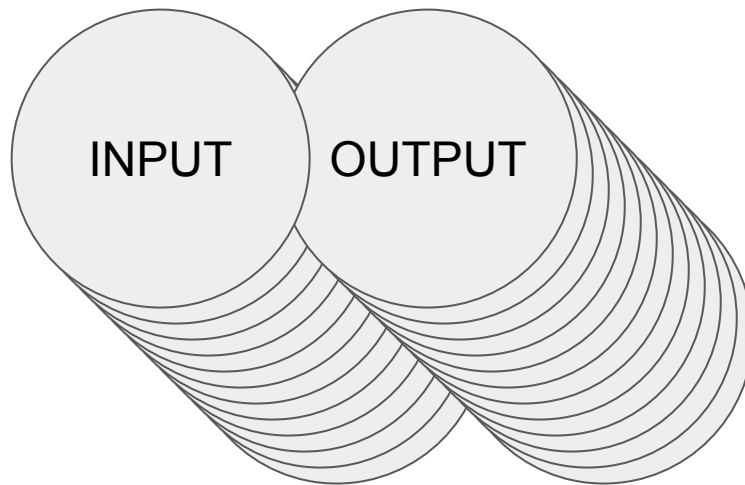
**EXPERIENCE**

**TASK ✓**

**PERFORMANCE  
MEASURE**

# Experiencia: Dataset

Un dataset es una colección de ejemplos de la tarea que se quiere aprender.



Cada ejemplo es una colección de mediciones.

Usualmente vectores



# Design Matrix

Object	Weight (g)	Colour (0=Green, 1=Red)
Red Apple 1	147	0.90
Red Apple 2	159	0.70
Red Apple 3	170	0.77
Green Apple 1	163	0.17
Green Apple 2	151	0.13
Banana 1	104	0.10
Banana 2	119	0.15
Banana 3	113	0.34
Banana 4	122	0.23
Banana 5	125	0.30

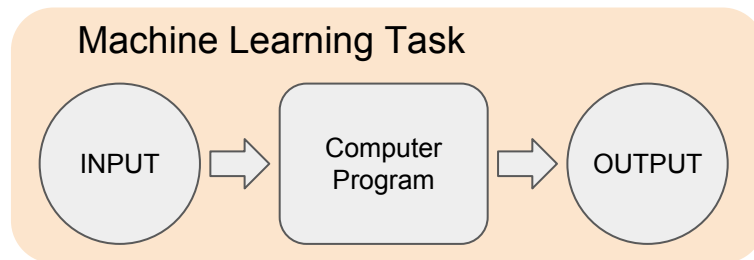


Data point

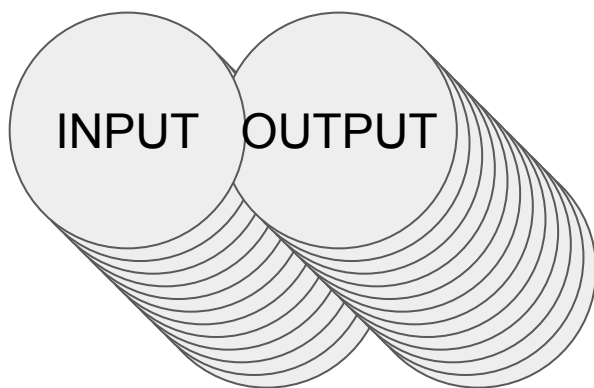


Lo que queremos aprender

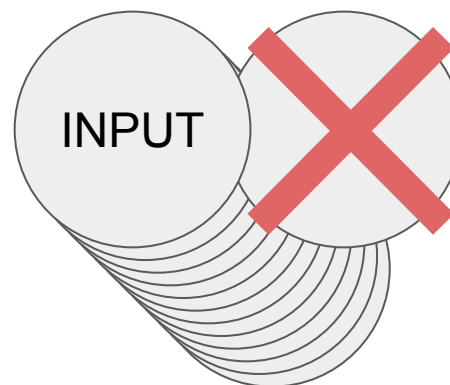
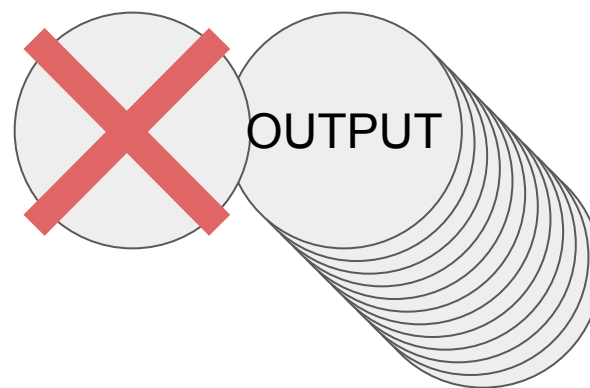
# Tipos de aprendizaje



SUPERVISADO



NO SUPERVISADO





# Datasets. Características deseables

- Que los ejemplos sean iid (independientes e idénticamente distribuidos)
- Que tengan la misma distribución que en el uso normal (posterior). Esto puede no ser trivial.
- Que tengan las mediciones relevantes para el problema, sin mediciones sin importancia.
- Suficientes datos



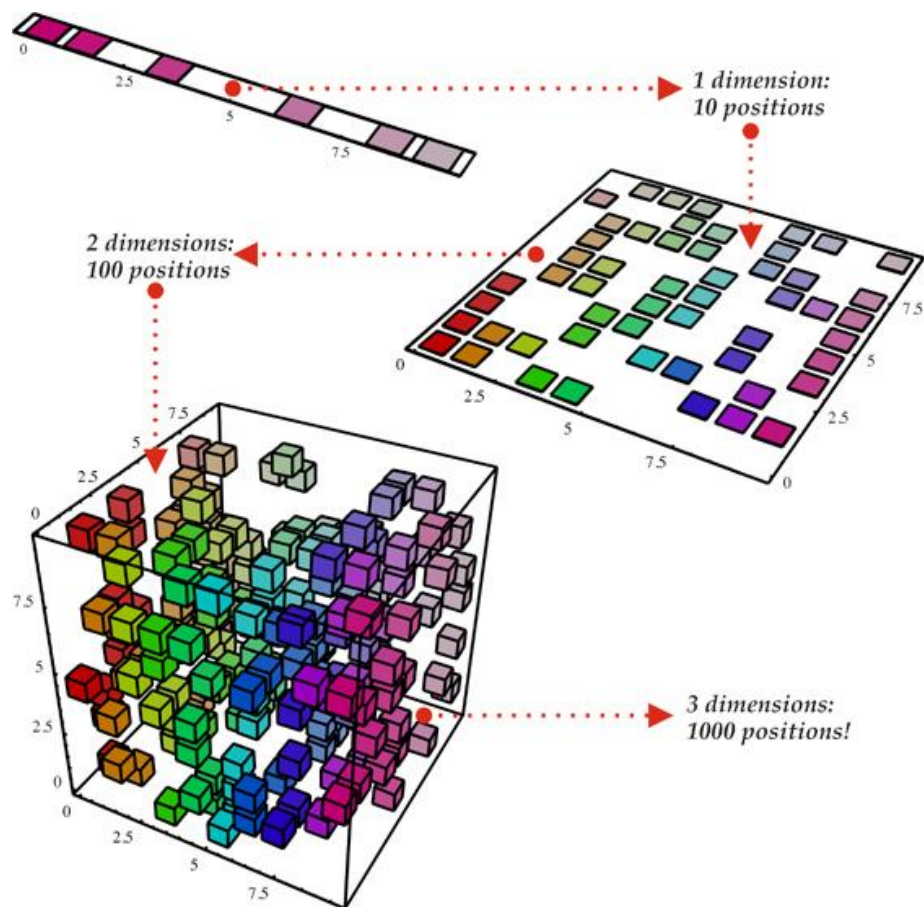


# Datasets. Posibles problemas

Qué pasa cuando no se cumplen estas cosas.

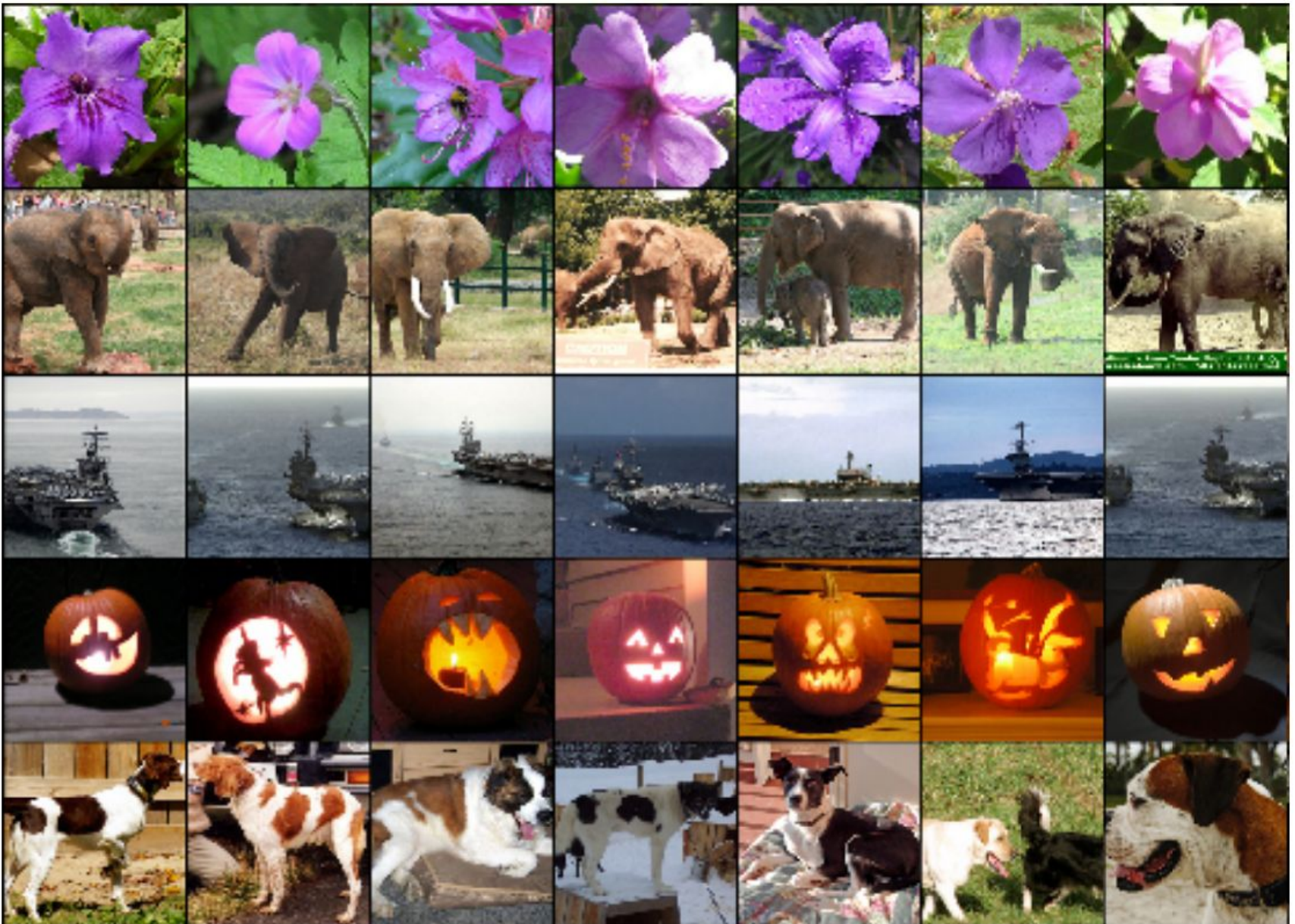
- Datos sesgados: comportamiento muy diferente en test.
- Pocos datos. Lo mismo. No se pueden capturar los detalles del sistema (real) que los genera.
- El sistema (real) genera datos con ruido. Se puede solucionar con más datos.

# Datasets. Problema de la alta dimensionalidad



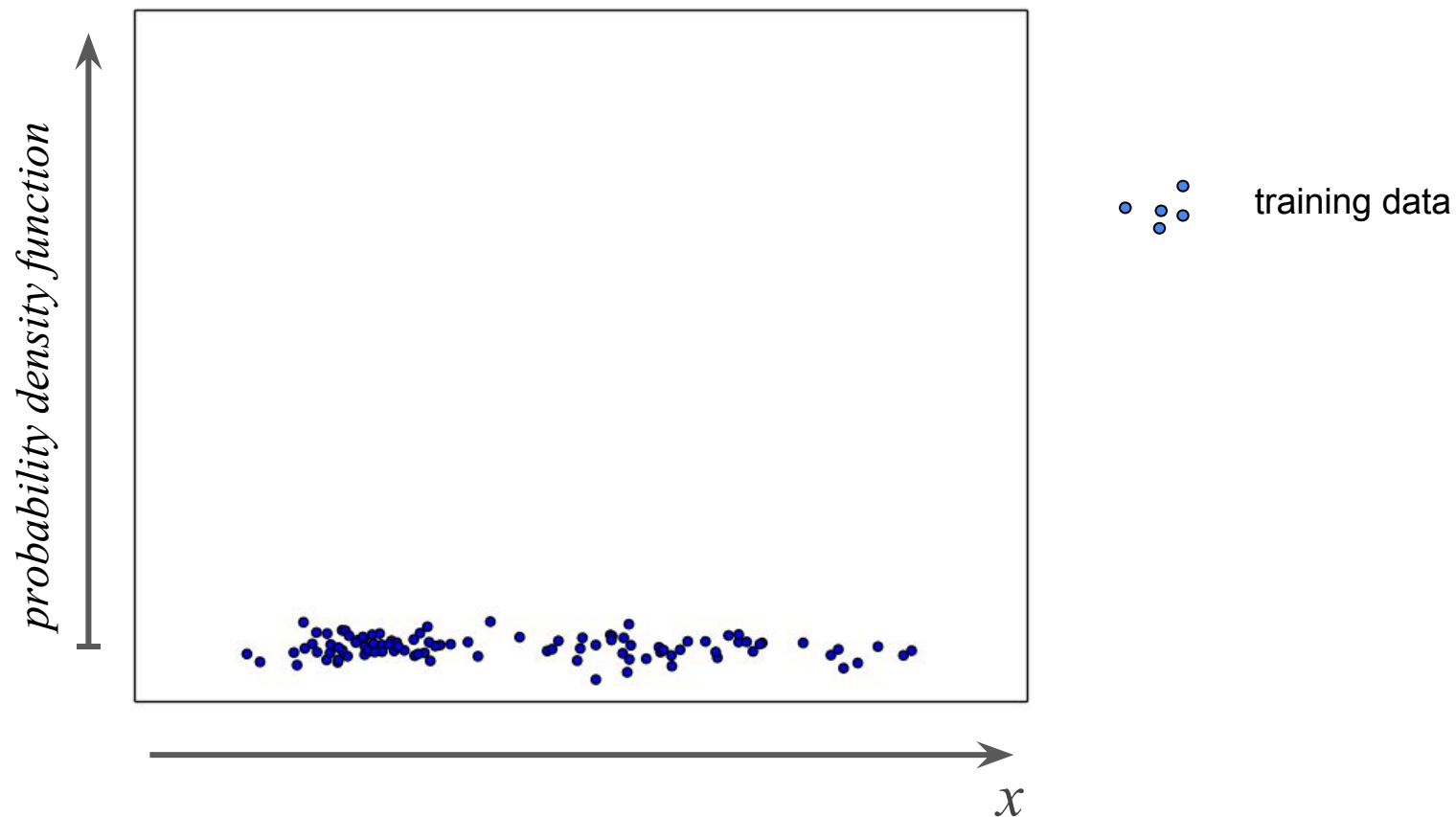
- La cantidad de datos necesaria para cubrir razonablemente bien un espacio crece exponencialmente con la cantidad de dimensiones.
- Las distancias (euclideas) se empiezan a comportar de manera contraintuitiva

Beyer, Kevin, et al. "When is "nearest neighbor" meaningful?." *International conference on database theory*. Springer Berlin Heidelberg, 1999.

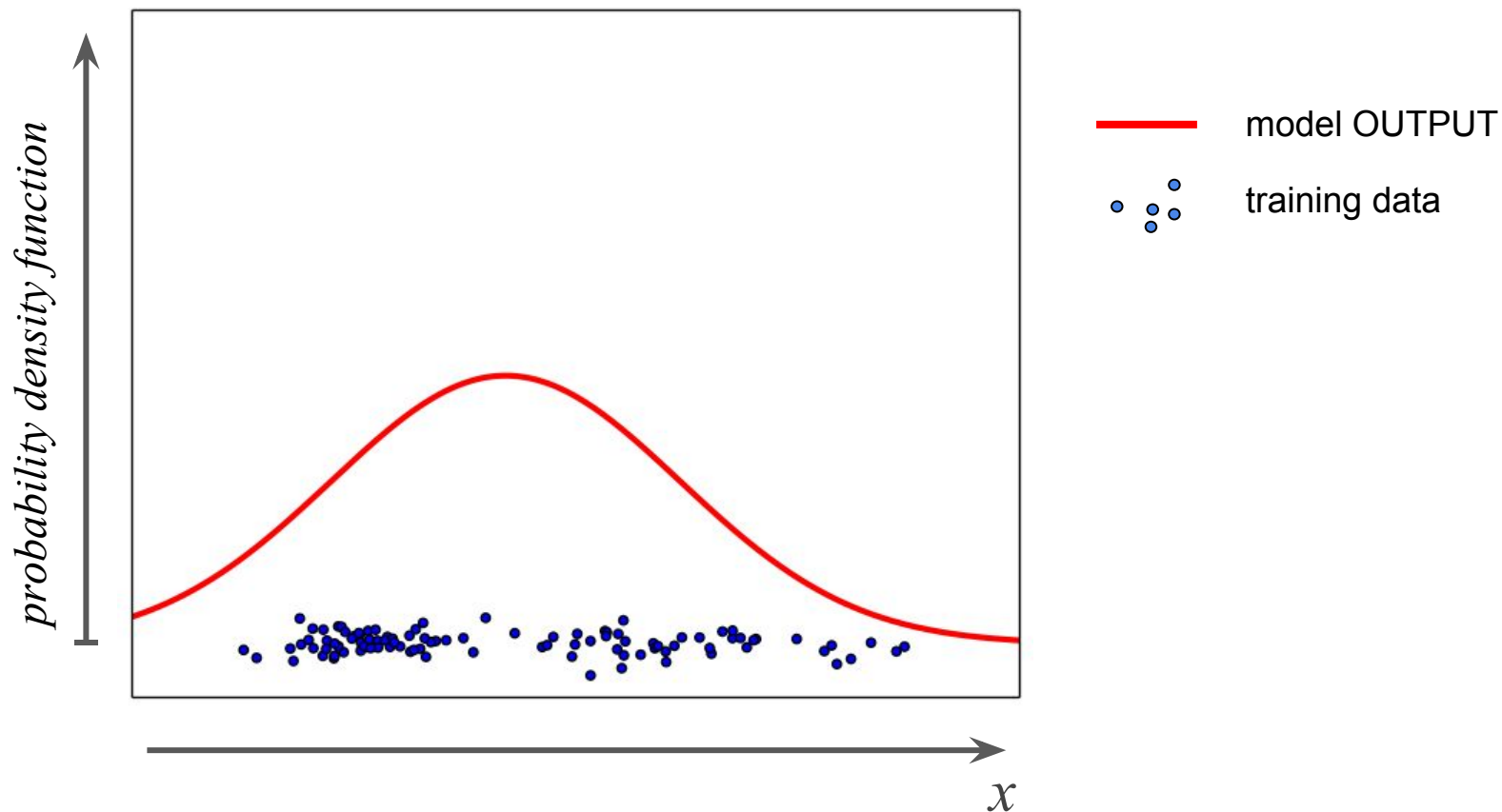


Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.

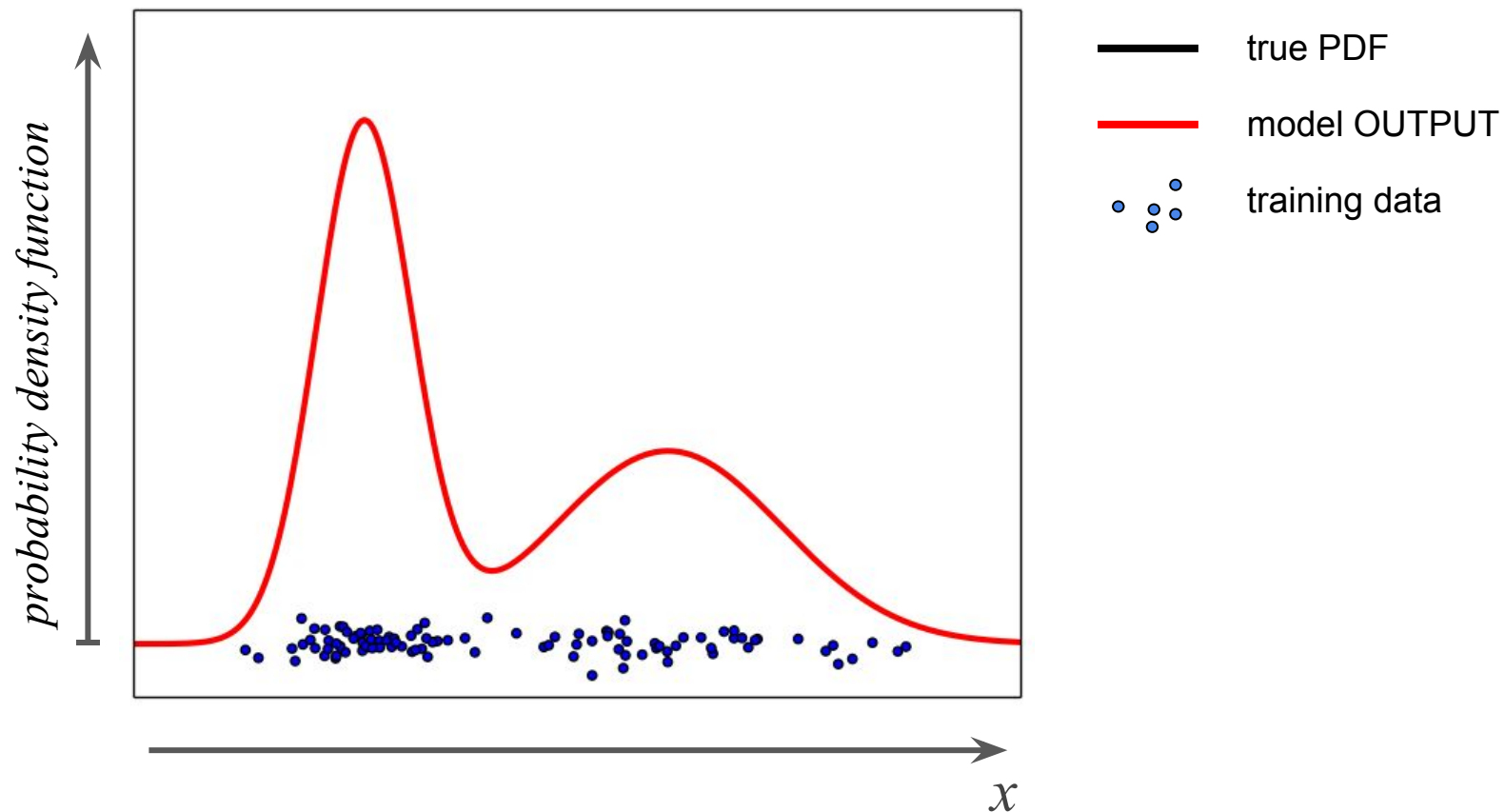
# Ejemplo: Aproximación de una PDF



# Modelo simple: una distribución normal.

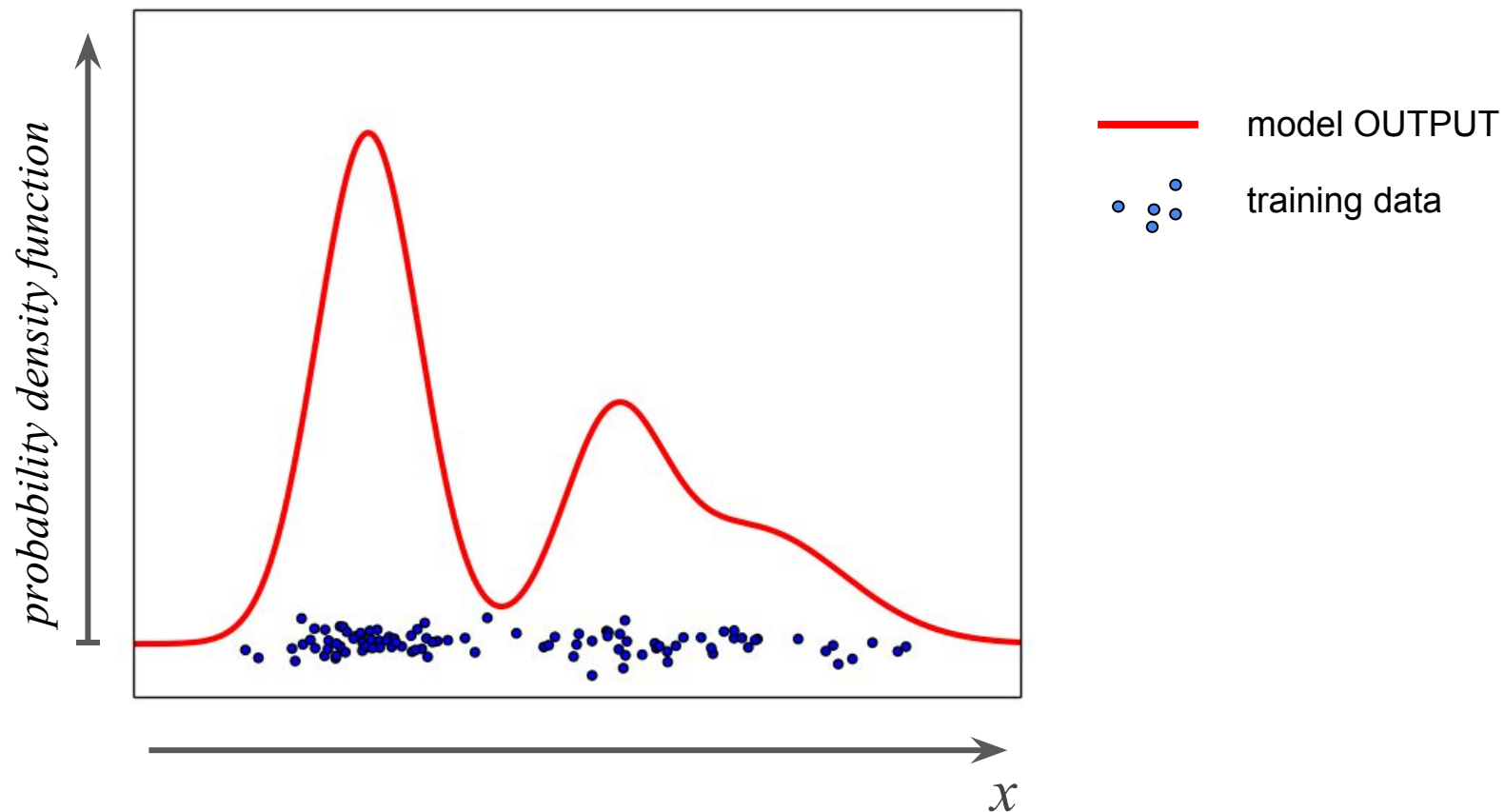


# Modelo más complejo: 2 gaussianas

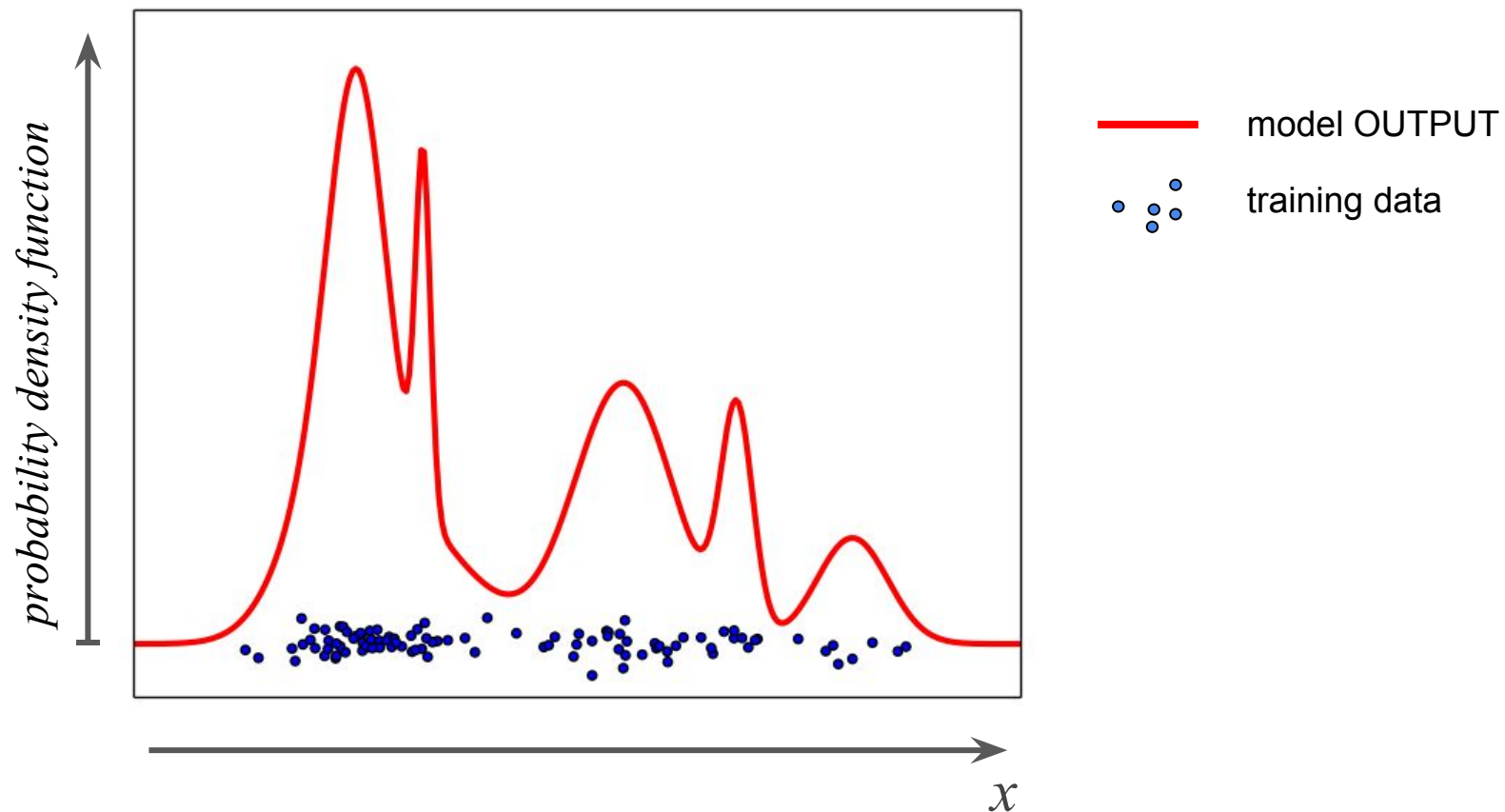




Más todavía: 3 gaussianas.

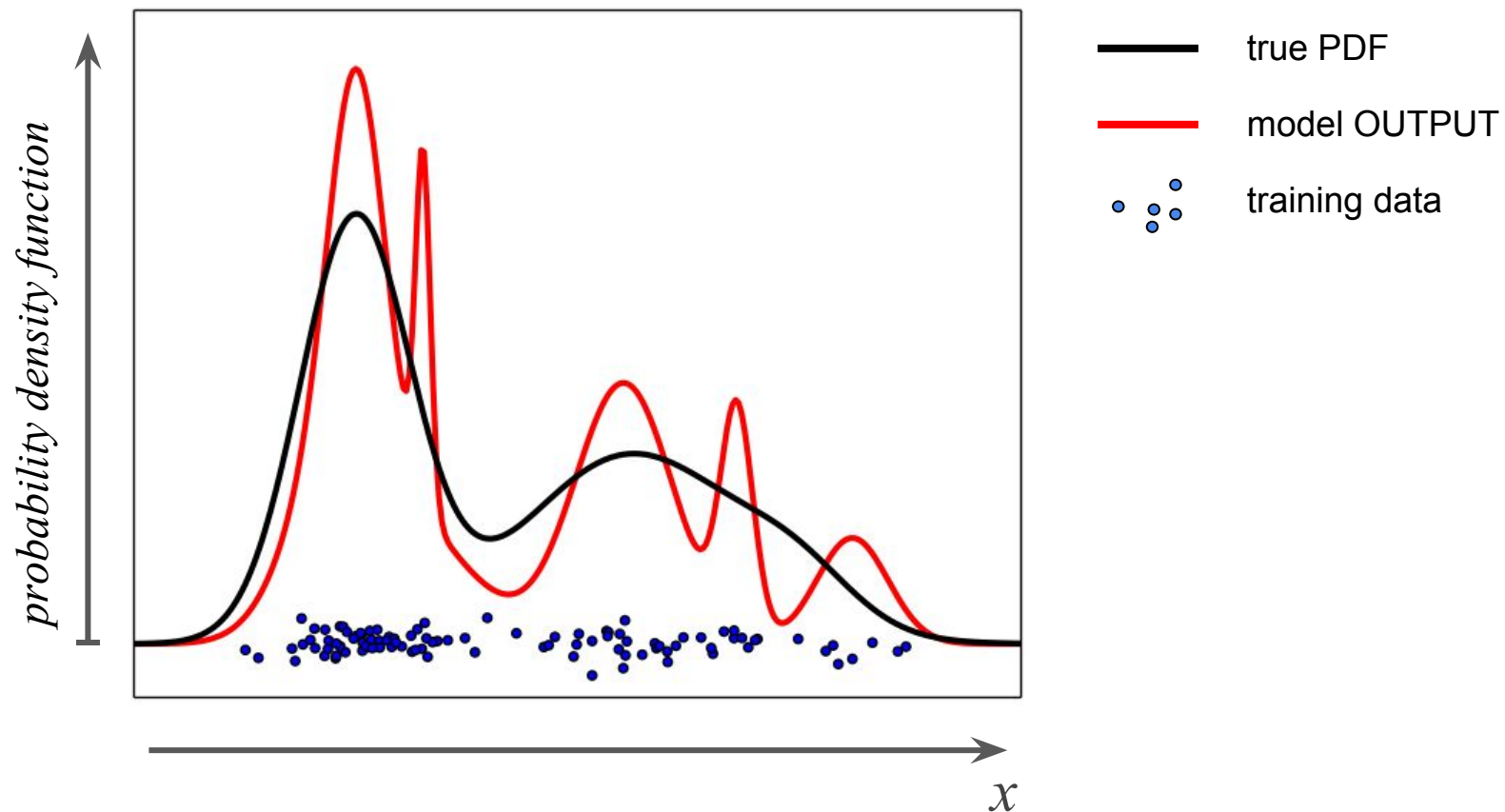


# 10 gaussianas

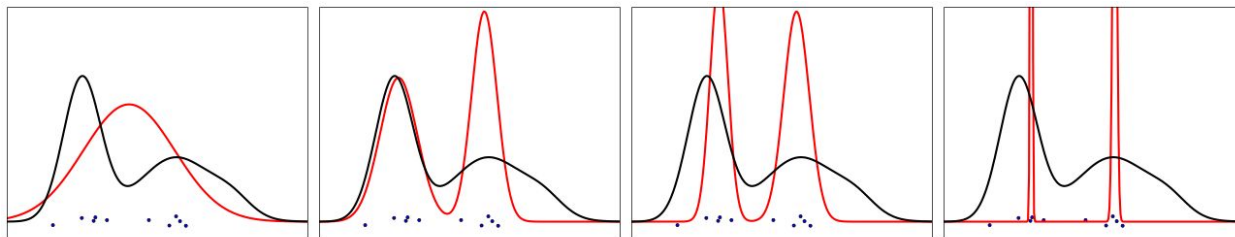




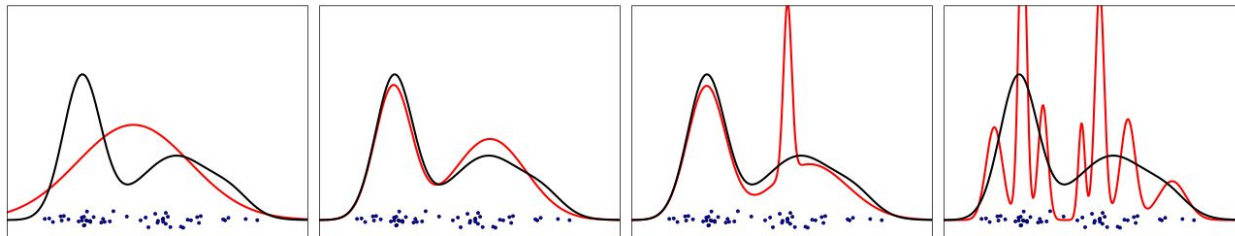
# 10 gaussianas: Overfitting



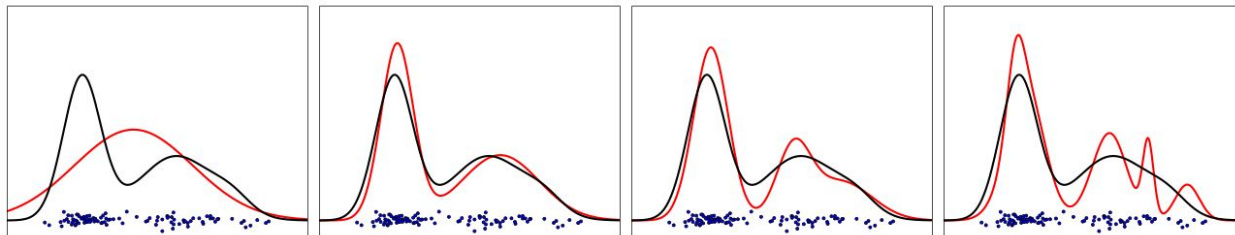
$N = 10$



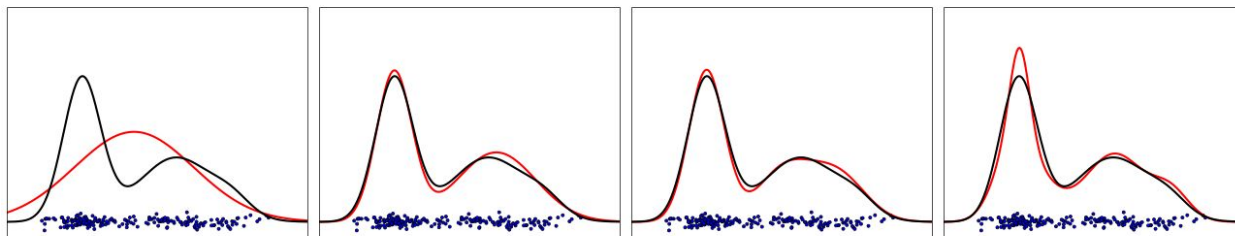
$N = 50$



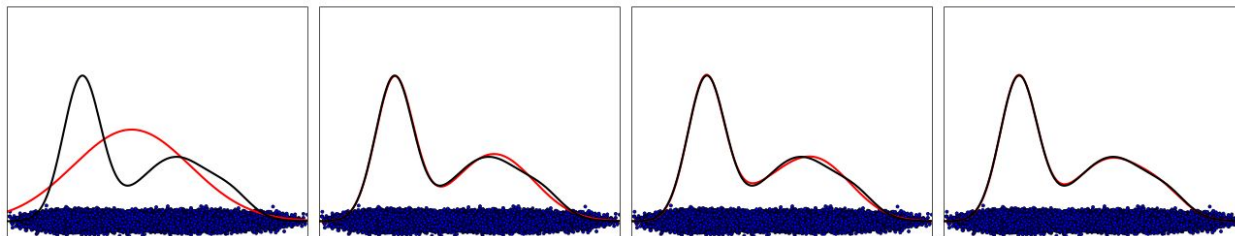
$N = 100$



$N = 200$



$N = 10^5$



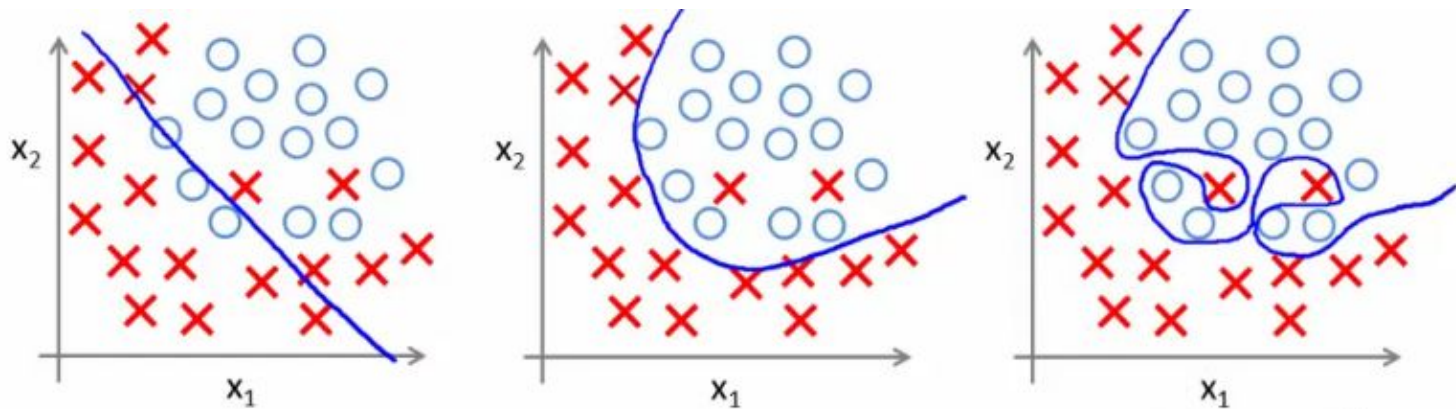
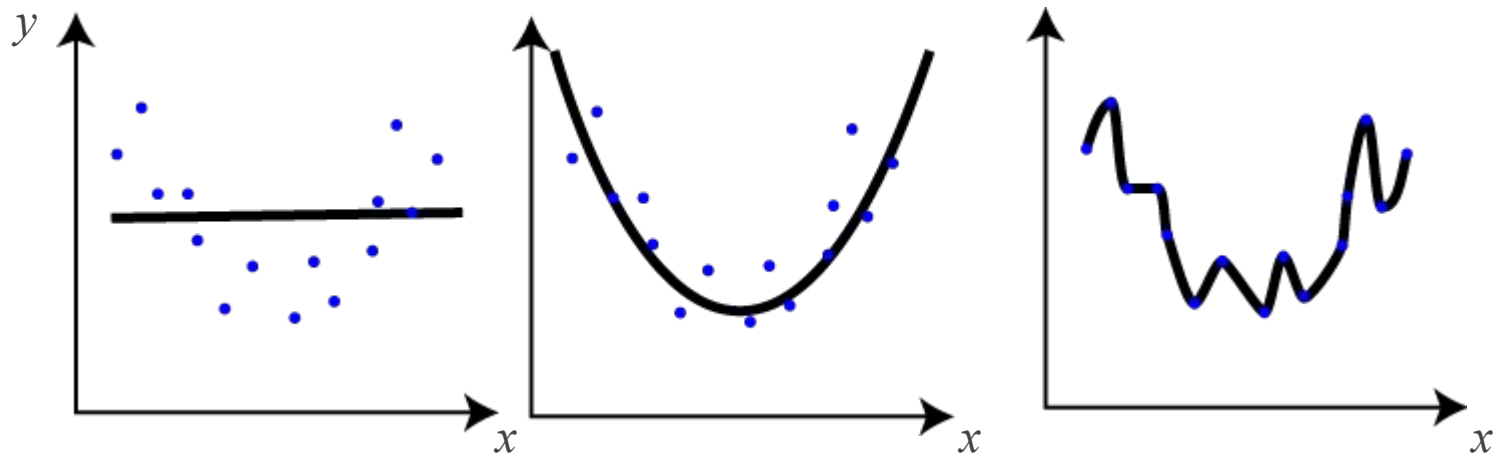
$G = 1$

$G = 2$

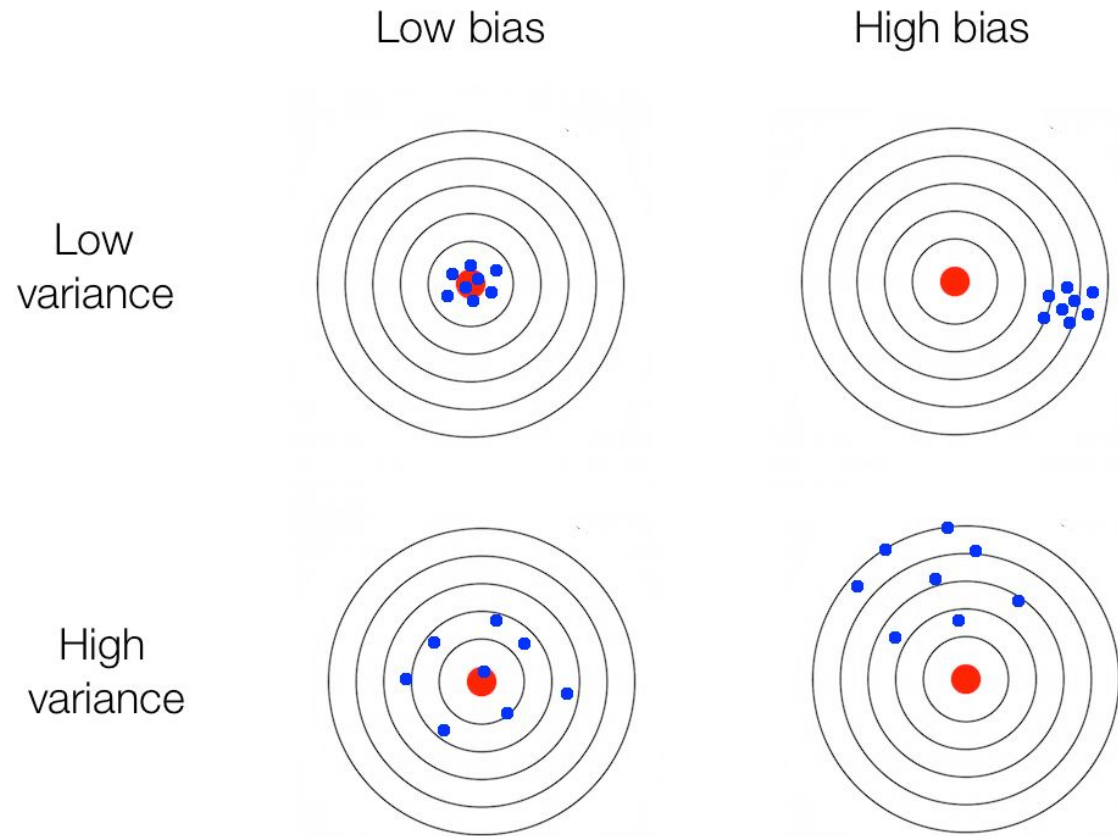
$G = 3$

$G = 10$

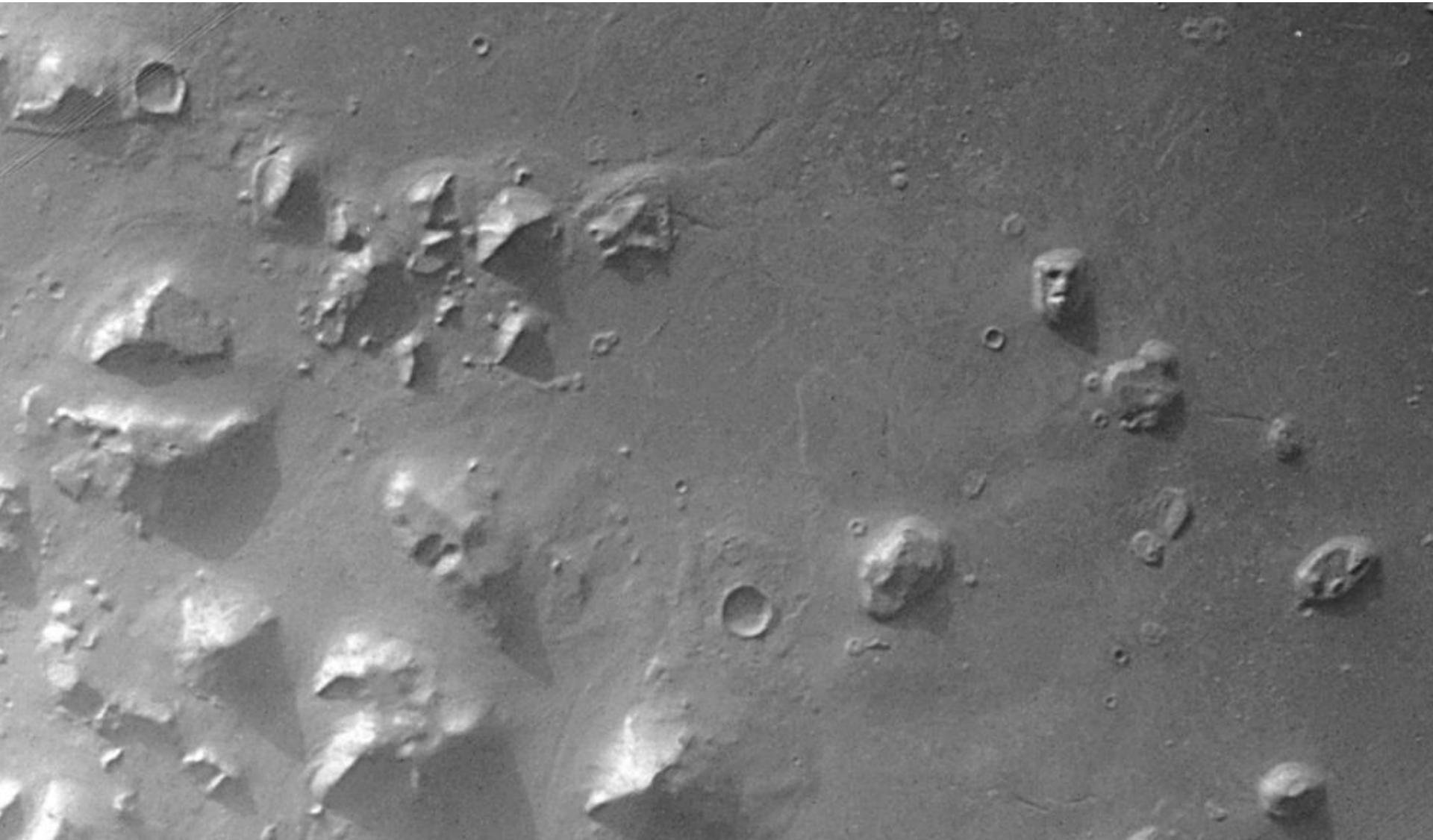
# Underfitting y overfitting



# Bias - Variance



Capacidad y subreajuste.



# Capacidad y sobreajuste

