ورشة عمل عن أساسيات الذكاء الاصطناعي وتطبيقاته

تقديم
م. لؤي بن زقطة

محاور ورشة العمل

01

مقدمة عن الذكاء
الاصطناعي وفروعه

02

أساسيات التعلم الآلي
Machine Learning

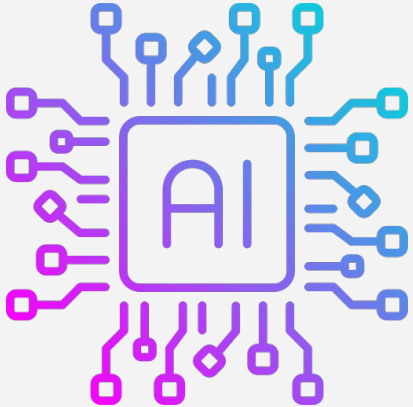
03

أساسيات التعلم العميق
Deep Learning

01

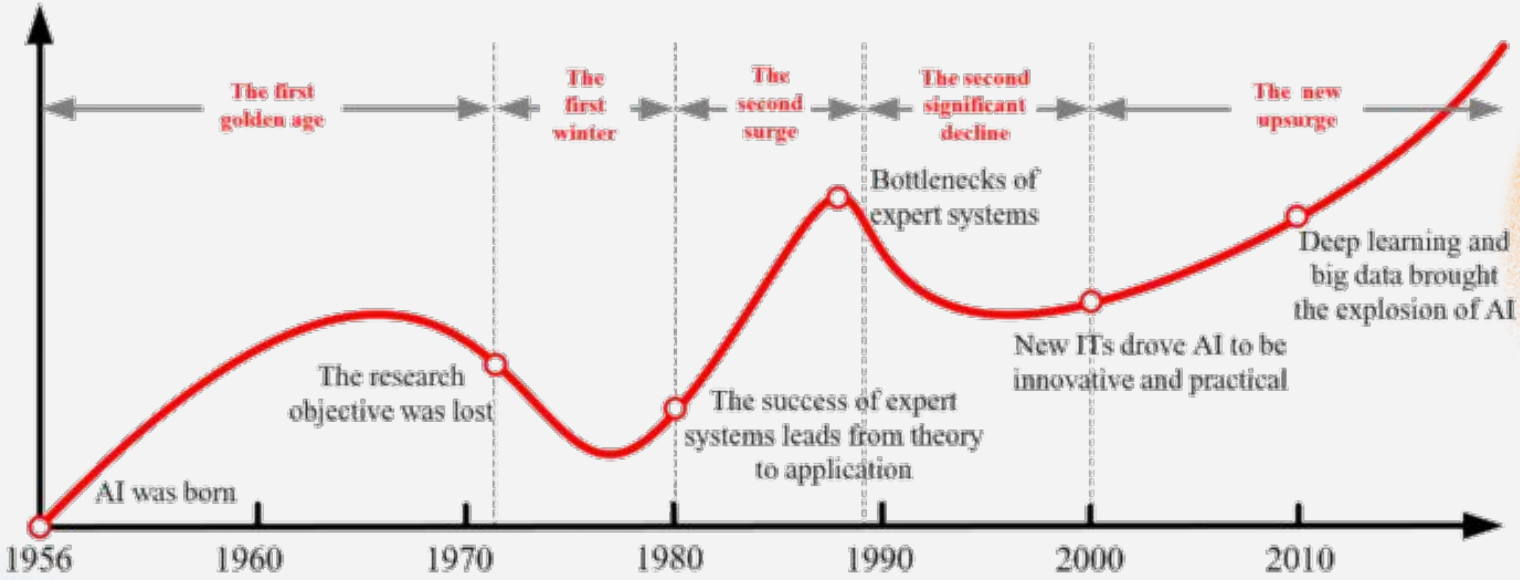
مقدمة عن الذكاء الاصطناعي وفروعه

ما هو الذكاء الاصطناعي؟



الذكاء الاصطناعي هو مجموعة من التقنيات التي تمكن أجهزة الكمبيوتر من أداء مجموعة متنوعة من الوظائف المتقدمة، بما في ذلك القدرة على رؤية وفهم وترجمة اللغة المنطوقة والمكتوبة، وتحليل البيانات، وتقديم التوصيات، والمزيد...

تاريخ الذكاء الاصطناعي



Geoffrey Hinton

FOLLOW

GET MY OWN PROFILE

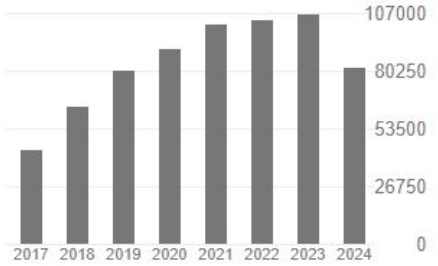
Emeritus Prof. Computer Science, University of Toronto
Verified email at cs.toronto.edu - Homepage

machine learning psychology artificial intelligence cognitive science computer science

TITLE	CITED BY	YEAR
Imagenet classification with deep convolutional neural networks A Krizhevsky, I Sutskever, GE Hinton Advances in neural information processing systems 25	164261 *	2012
Deep learning Y LeCun, Y Bengio, G Hinton Nature 521 (7553), 436-44	85739	2015
Learning internal representations by error-propagation DE Rumelhart, GE Hinton, RJ Williams Parallel Distributed Processing: Explorations in the Microstructure of ...	55034 *	1986
Dropout: a simple way to prevent neural networks from overfitting N Srivastava, G Hinton, A Krizhevsky, I Sutskever, R Salakhutdinov The journal of machine learning research 15 (1), 1929-1958	53229	2014

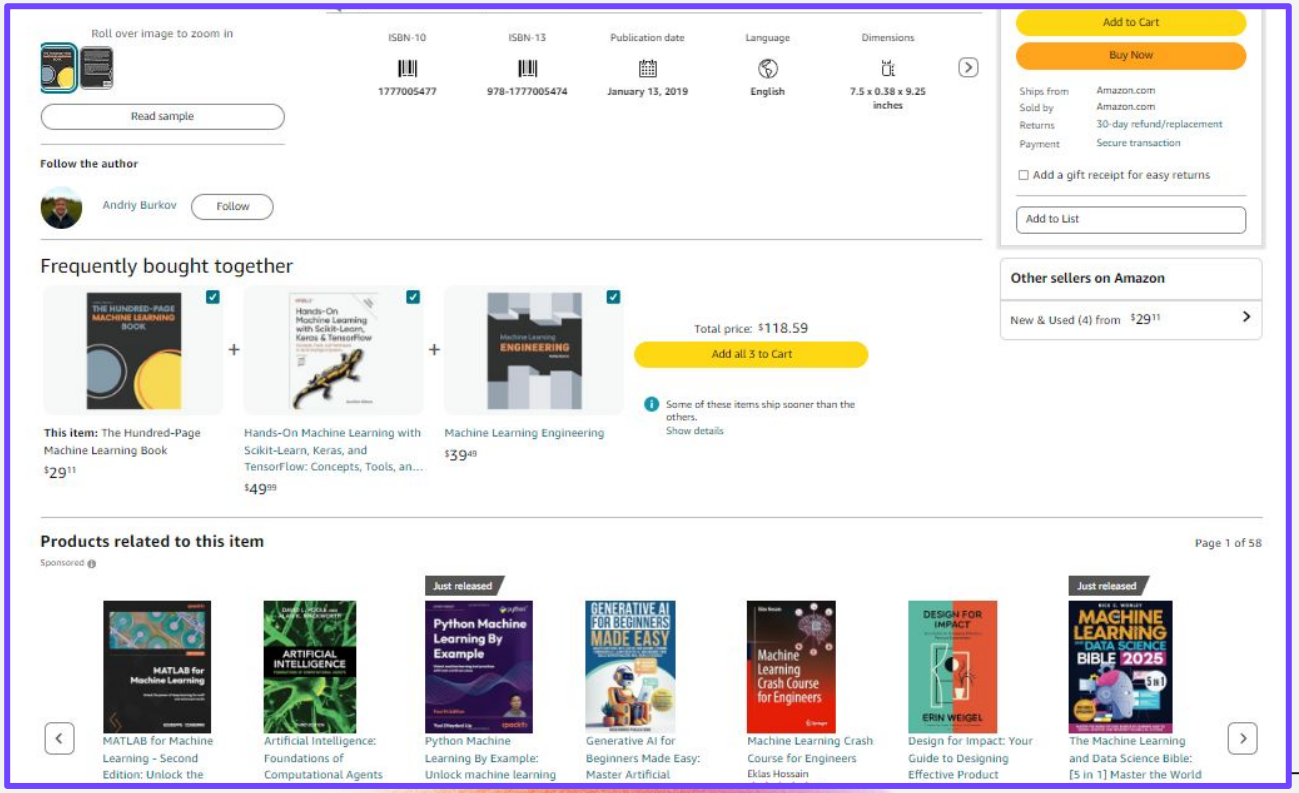
Cited by VIEW ALL

	All	Since 2019
Citations	859619	565700
h-index	187	137
i10-index	482	366



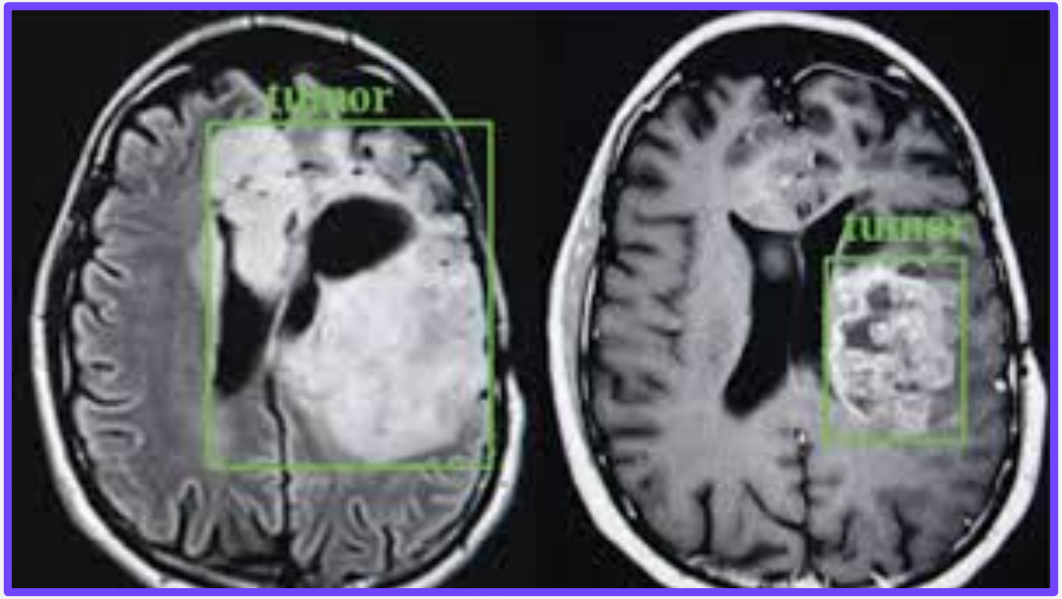
تطبيقات
الذكاء الاصطناعي

أنظمة التوصية
Recommendation Systems



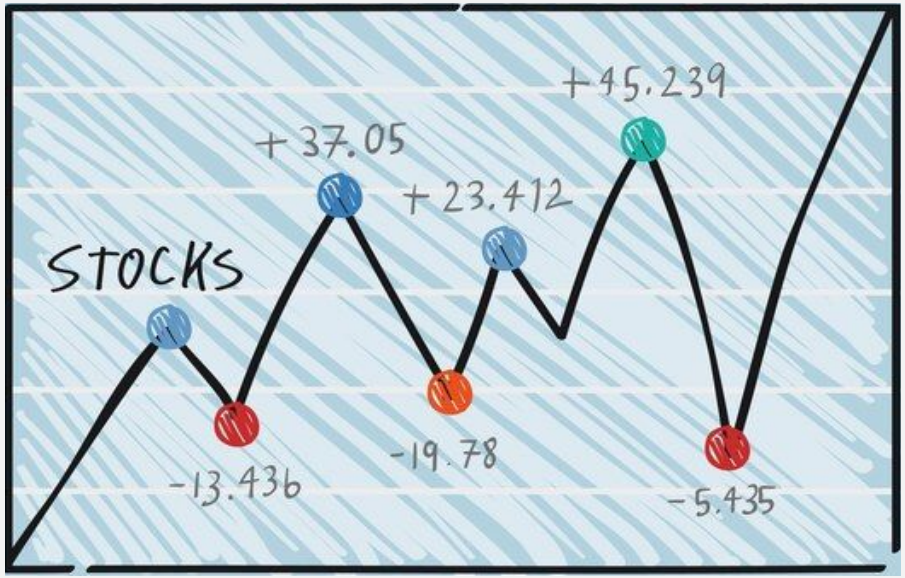
التعرف على الكلام
Speech Recognition

الكشف عن الأورام
Tumor detection



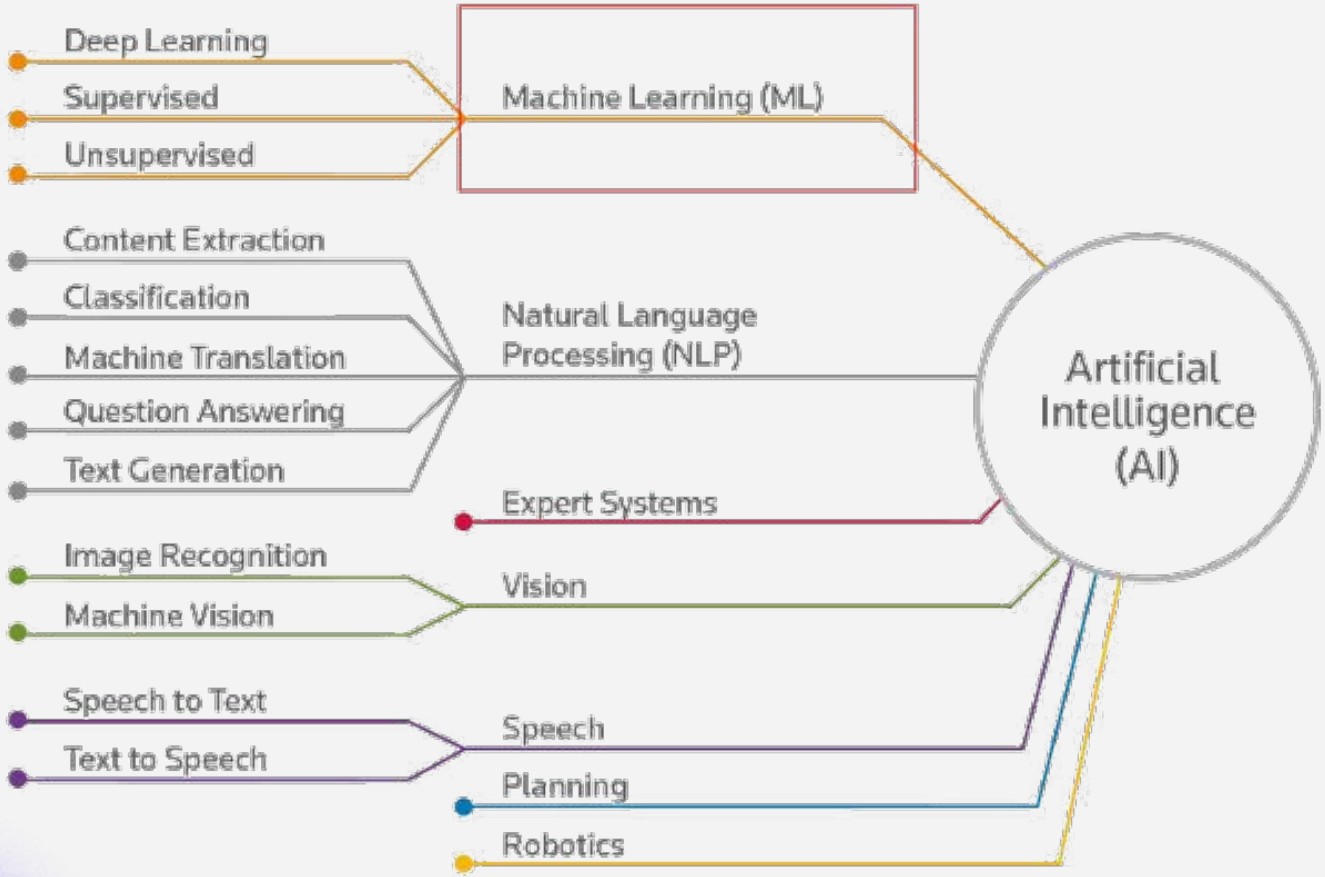
القيادة الذاتية
Self-driving

التنبؤ بالأسهم
Stock prediction



أنواع الذكاء الاصطناعي

- الذكاء الاصطناعي الضيق
  Narrow AI
- الذكاء الاصطناعي العام
  General AI


Deep Learning
Supervised
Unsupervised
Machine Learning (ML)
Content Extraction
Classification
Machine Translation
Question Answering
Text Generation
Natural Language Processing (NLP)
Artificial Intelligence (AI)
Expert Systems
Image Recognition
Machine Vision
Vision
Speech to Text
Text to Speech
Speech
Planning
Robotics

02

أساسيات التعلم الآلي
Machine Learning

كيف نتعلم ؟

من خلال رؤية الأمثلة،
العديد والعديد منها .....

كيف نتعلم ؟

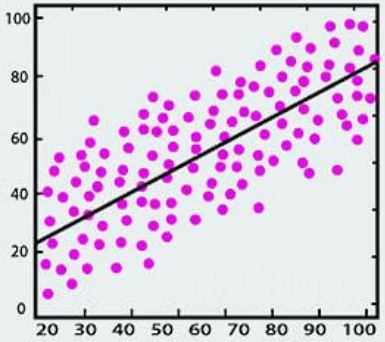
المحاولة والخطأ
والاستكشاف ...

أنواع التعلم الآلي (ML)

التعلم الخاضع للإشراف
Supervised Learning

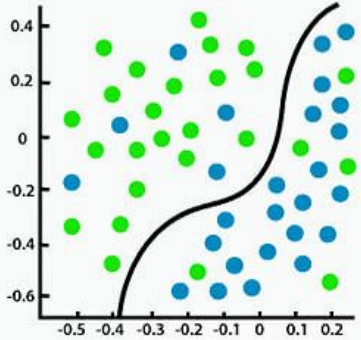
التعلم غير الخاضع للإشراف
Unsupervised Learning

التعلم الخاضع للإشراف
supervised learning



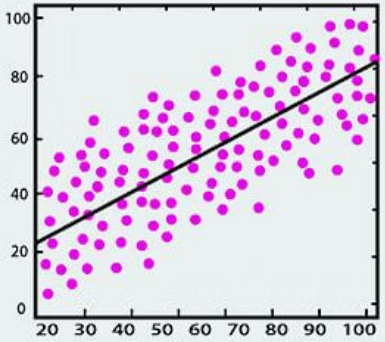
Regression

versus



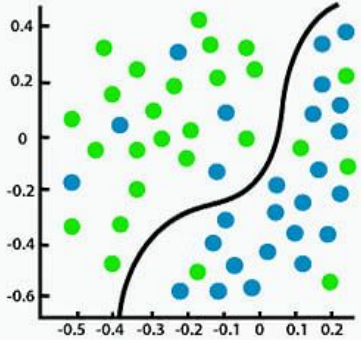
Classification

التعلم الخاضع للإشراف
supervised learning

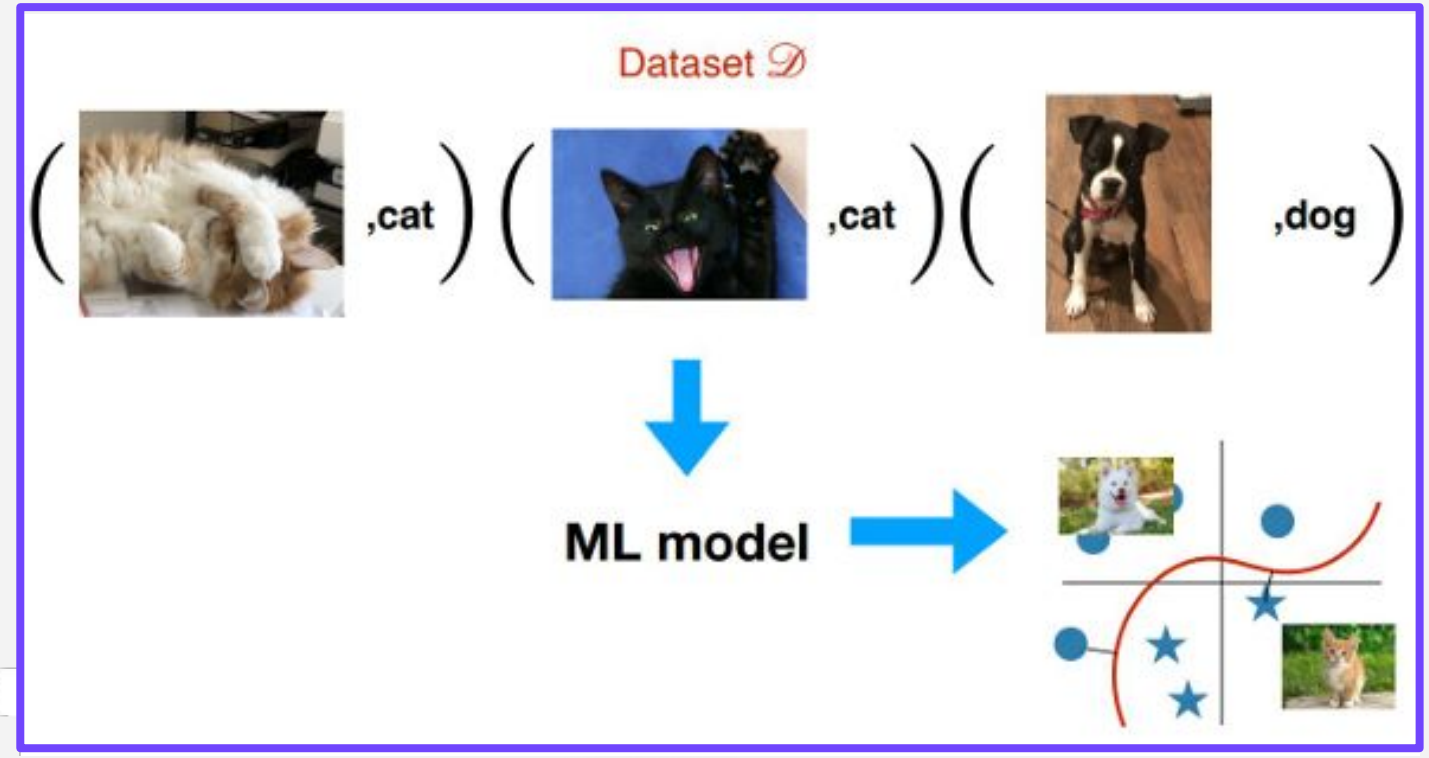


Regression

versus



Classification



Slide Credit: Cornell CS4780

الصياغة الرياضية

Dataset:

$$\mathcal{D} = \{(x_1, y_1), \ldots, (x_n, y_n)\}, x_i \in \mathbb{R}^d, y_i \in \mathcal{C}(\text{ e.g., } \mathcal{C} = \{-1, 1\}), (x_i, y_i) \sim \mathcal{P}$$

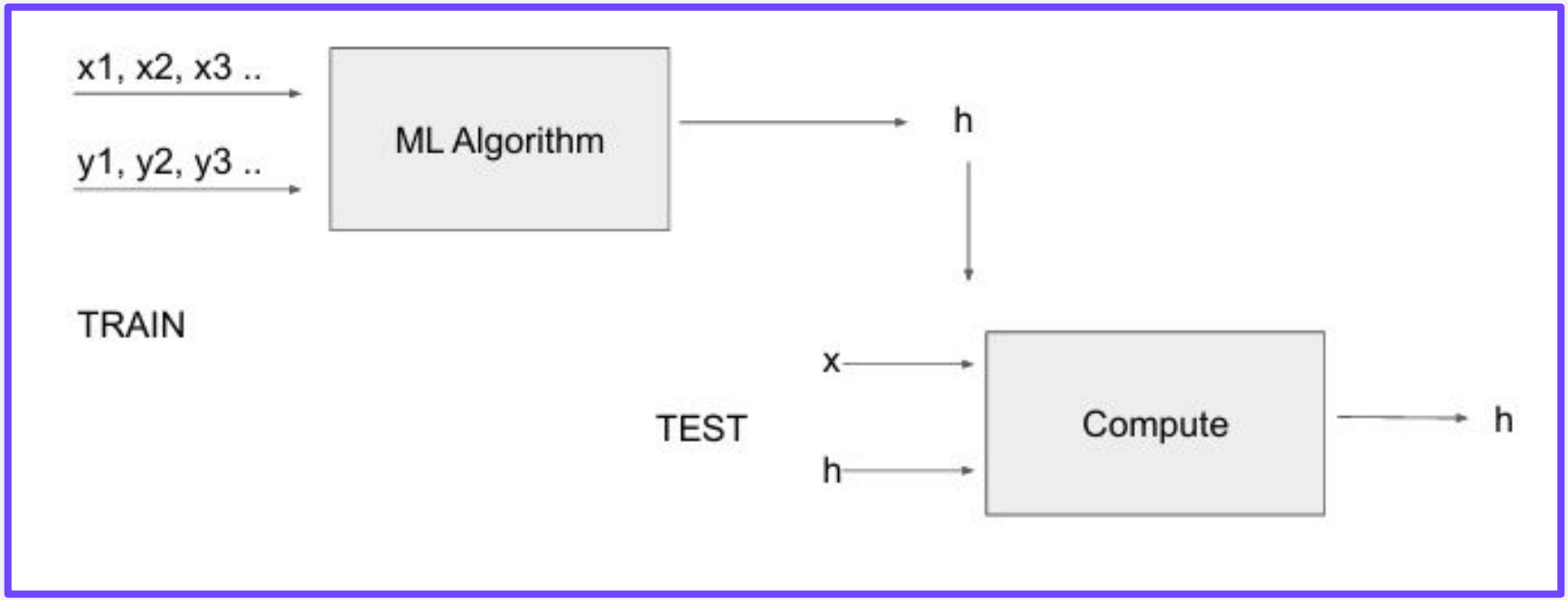
Hypothesis:

$$h : \mathbb{R}^d \mapsto \mathcal{C}$$

Hypothesis class

$$\mathcal{H} = \{h\}$$

Slide Credit: Cornell CS4780



Slide Credit: Cornell CS4780

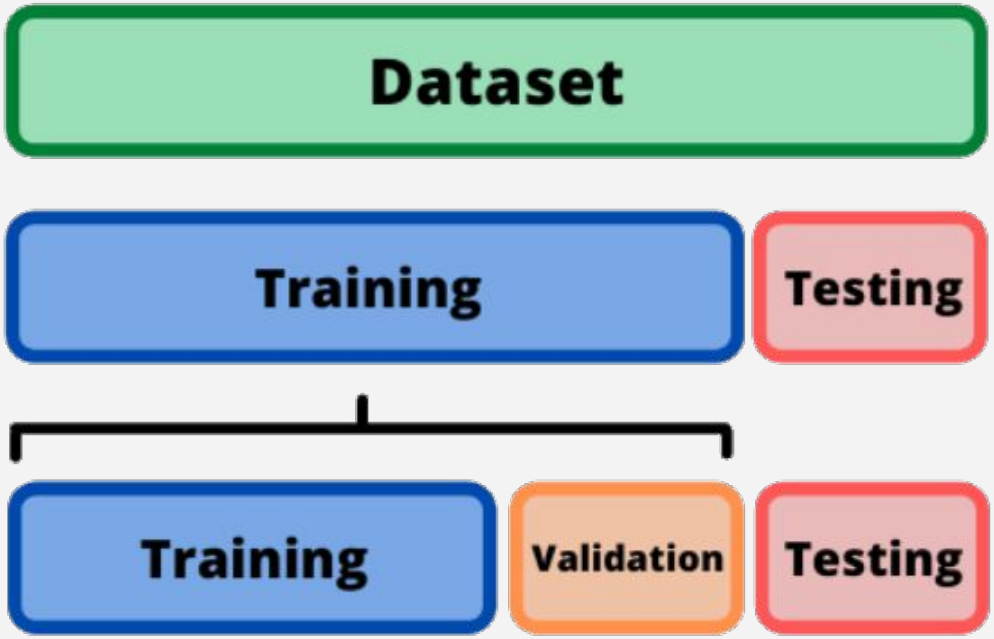
دالة الخسارة
The Loss Function

كيف يمكن اختيار أفضل خوارزمية h(.) ؟

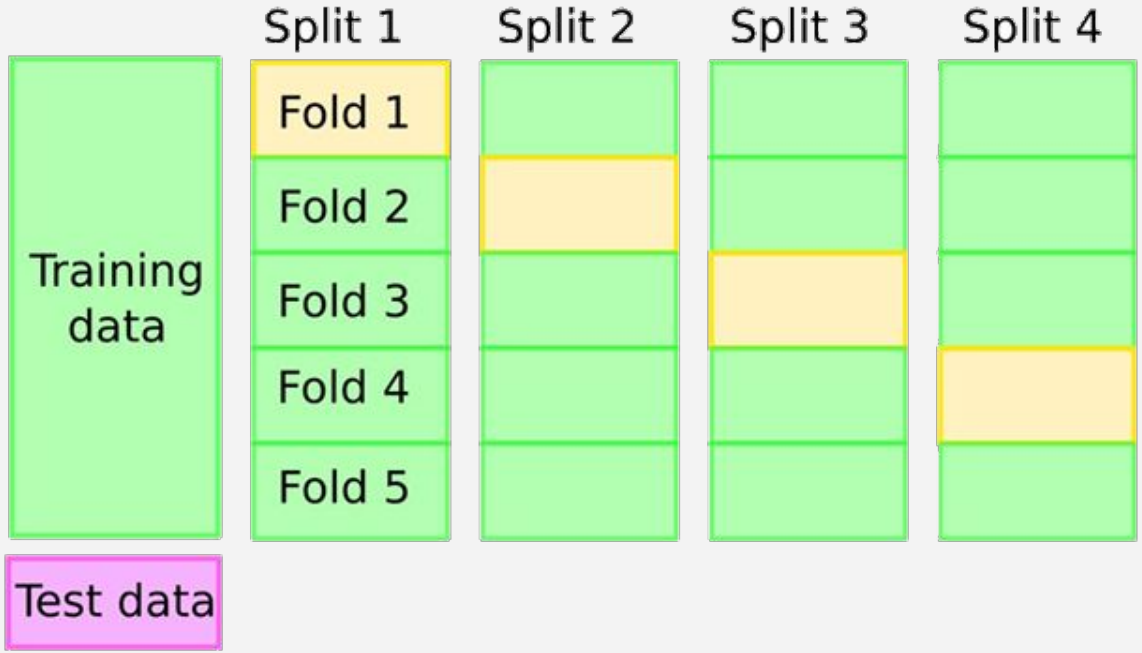
تقوم دالة الخسارة بتقييم الخوارزمية على
بيانات التدريب لدينا وتخبرنا بمدى سوءها

$$\mathcal{L}_{sq}(h) = \frac{1}{n}\sum_{i=1}^{n}(h(\mathbf{x}_i) - y_i)^2$$

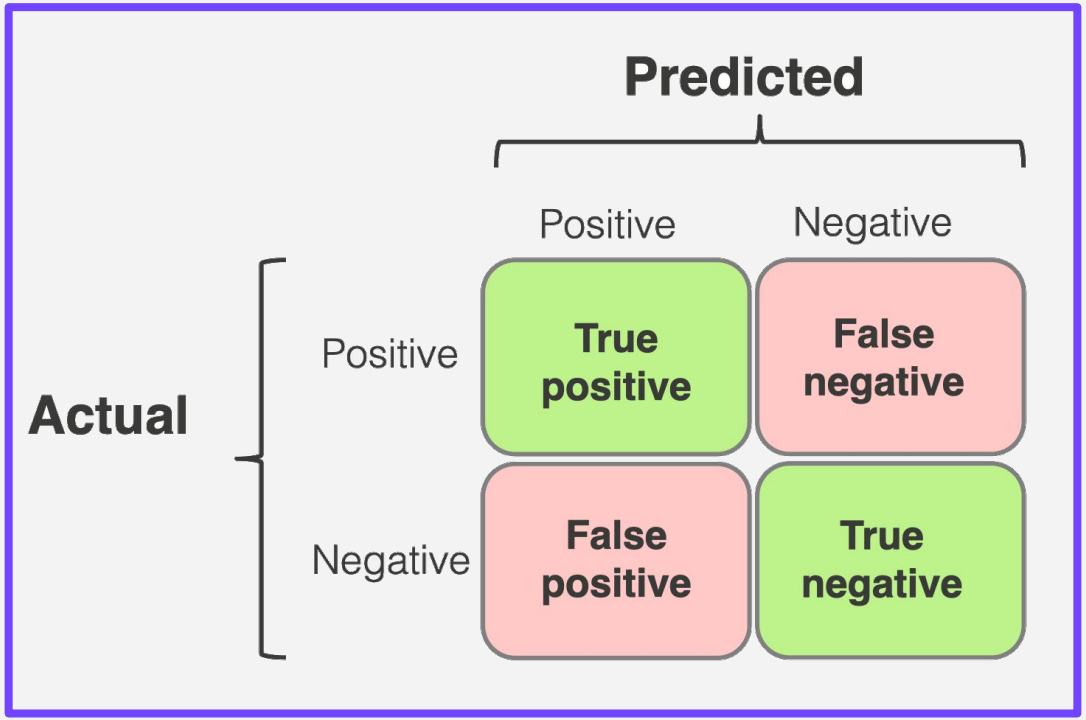
Slide Credit: Cornell CS4780


Dataset
Training
Testing
Training
Validation
Testing

Cross-Validation



Evaluation Metrics / مقاييس التقييم



Evaluation Metrics / مقاييس التقييم

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

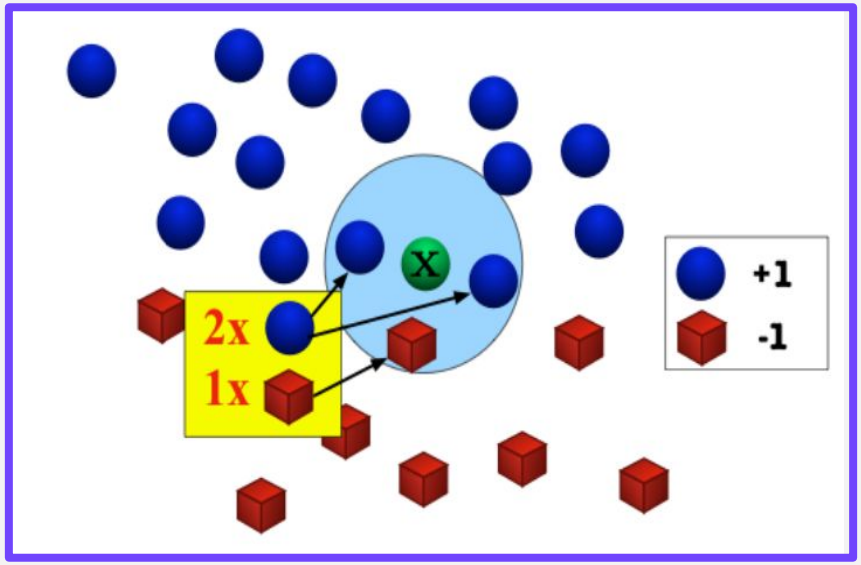
$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

خوارزمية 01

K-nearest neighbors (K-NN)

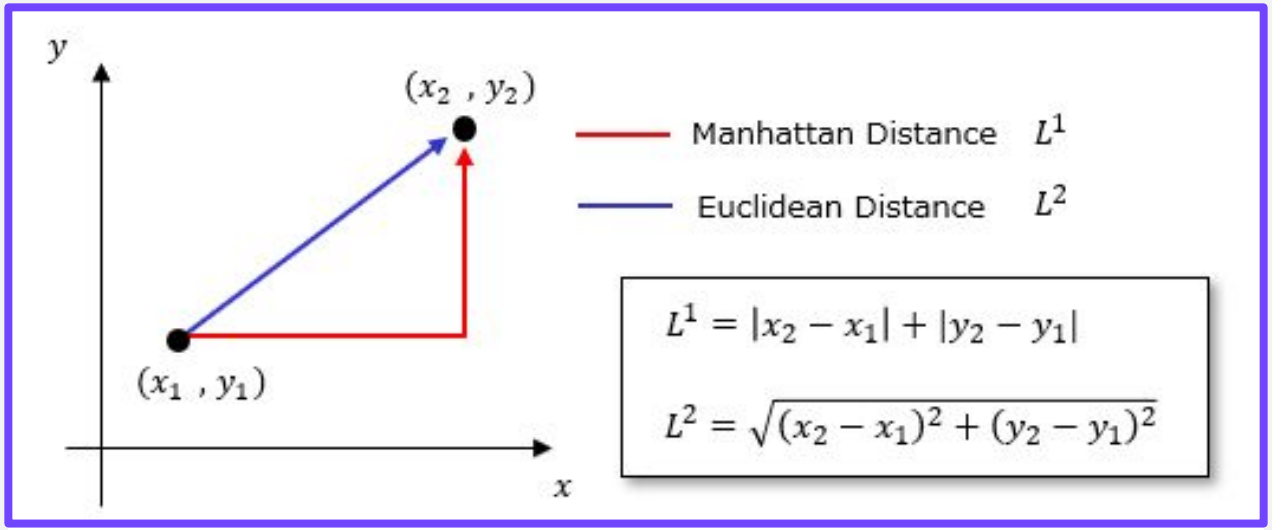
K-NN



الافتراض: المدخلات المتشابهة لها مخرجات متشابهة

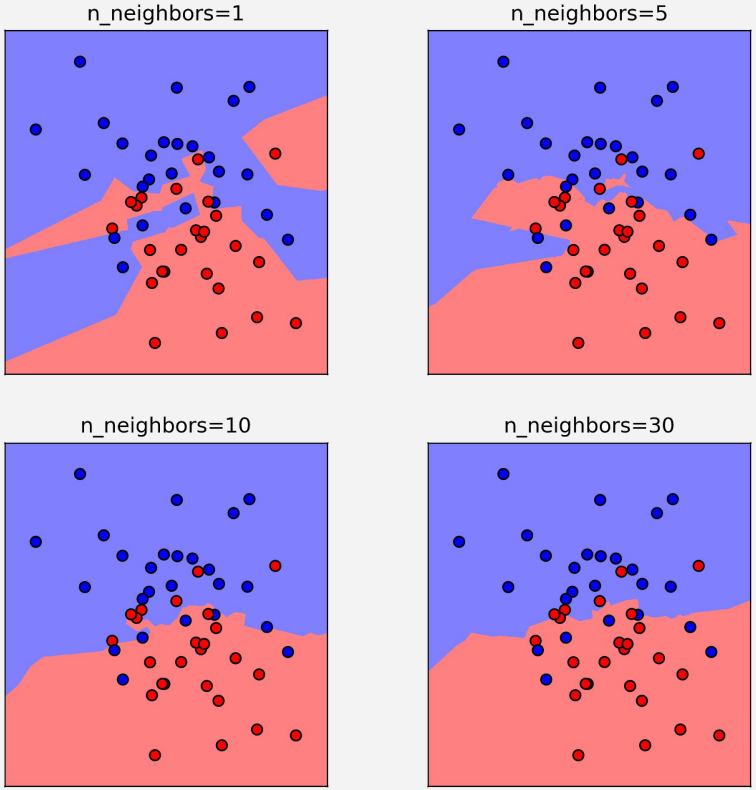
Slide Credit: Cornell CS4780

K-NN



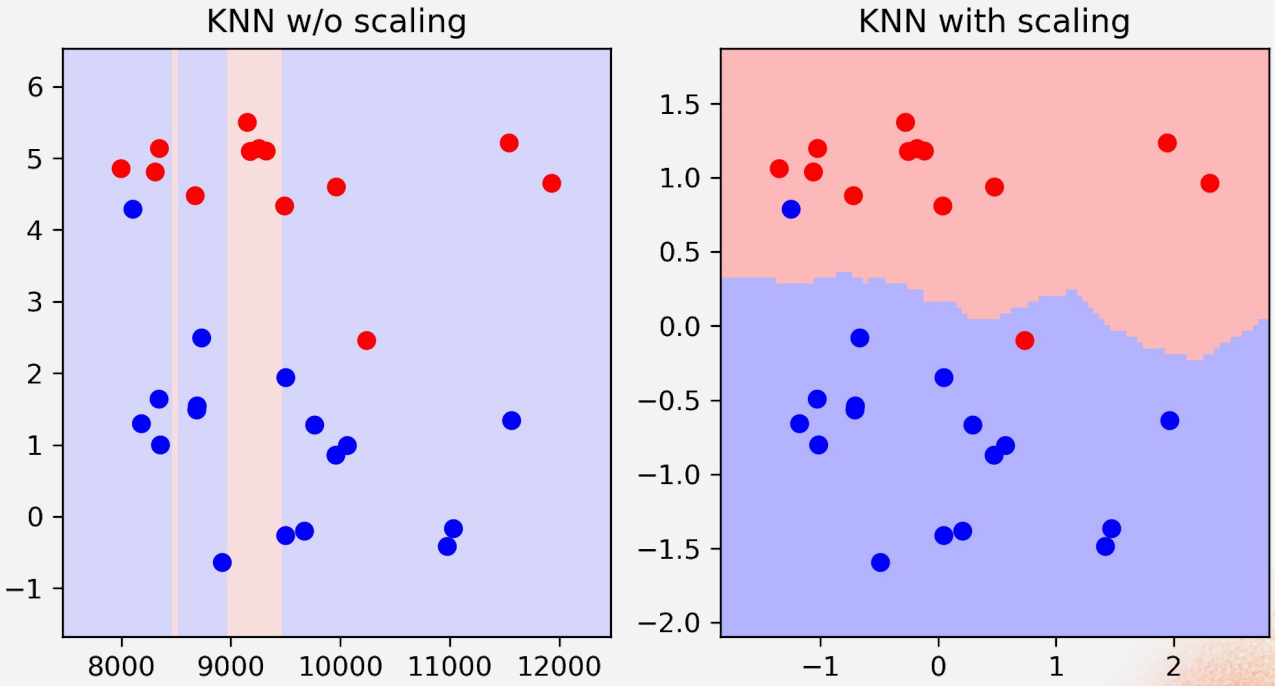
Slide Credit: Cornell CS4780

تـأثير قيمة K



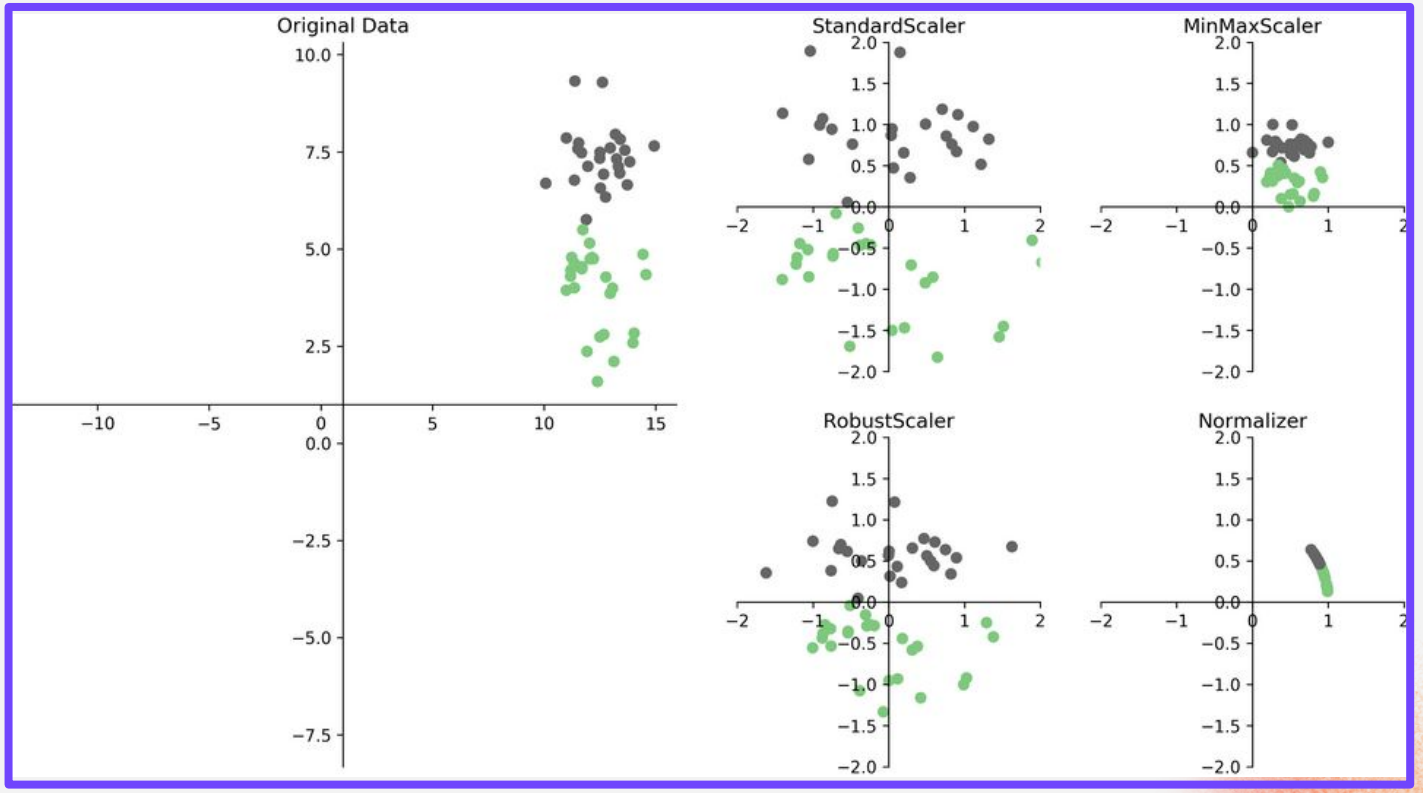
Slide Credit: Amuller, COMS4995

Feature Scaling



Slide Credit: Amuller, COMS4995

Feature Scaling



Slide Credit: Amuller, COMS4995

مزايا و عيوب KNN

بسيطة وبديهية

حدود قرار مرنة

يتكيف على الفور عندما
نجمع بيانات تدريب جديدة

التعقيد الحسابي

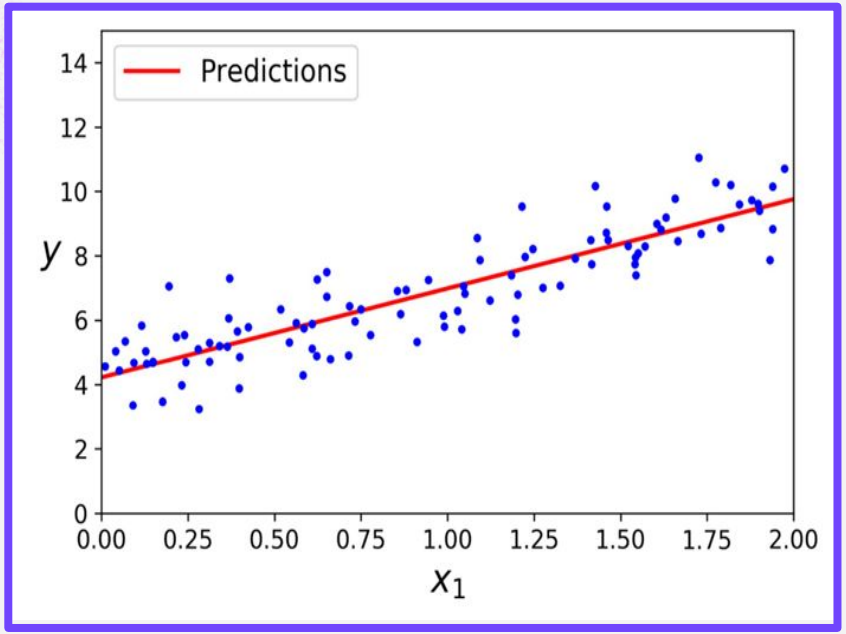
حساس للعينات غير ذات الصلة والمكررة

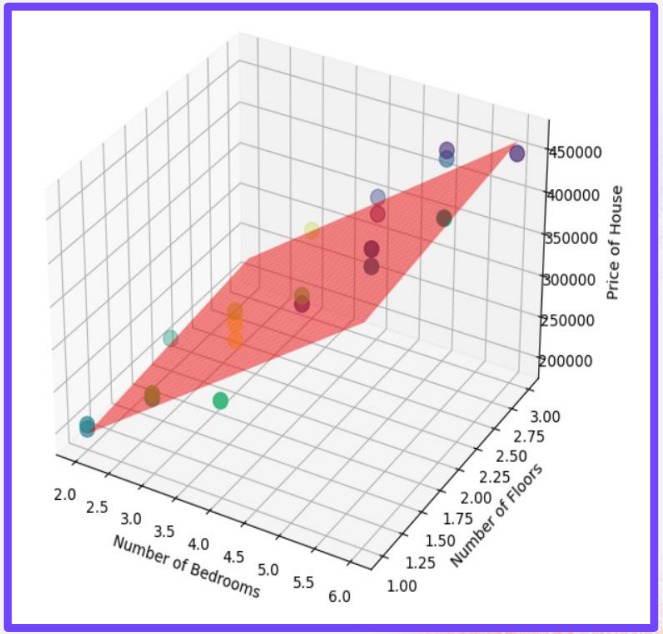
الأداء الضعيف مع البيانات عالية الأبعاد
(Curse of Dimensionality)

خوارزمية 02

Regression

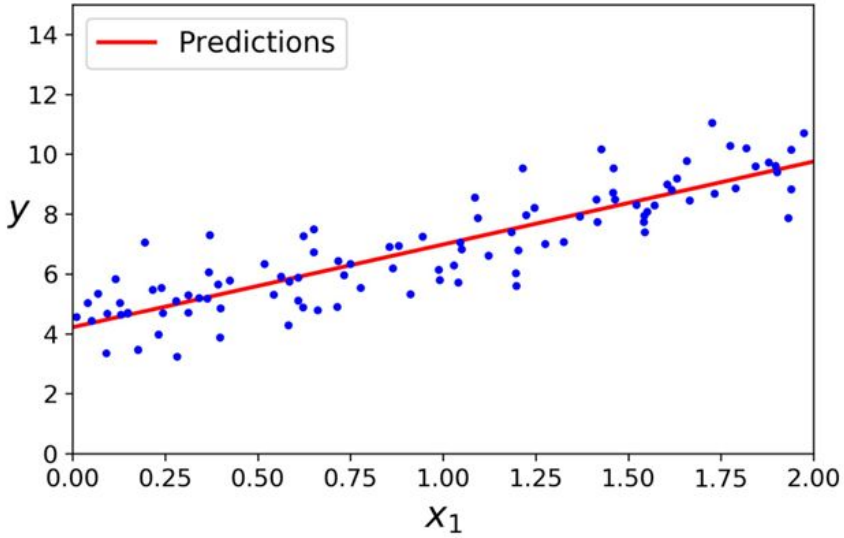
Linear Regression





Slide Credit: Amuller, COMS4995

الصياغة الرياضية



$$\hat{y} = w^T \mathbf{x} + b = \sum_{i=1}^{p} w_i x_i + b$$

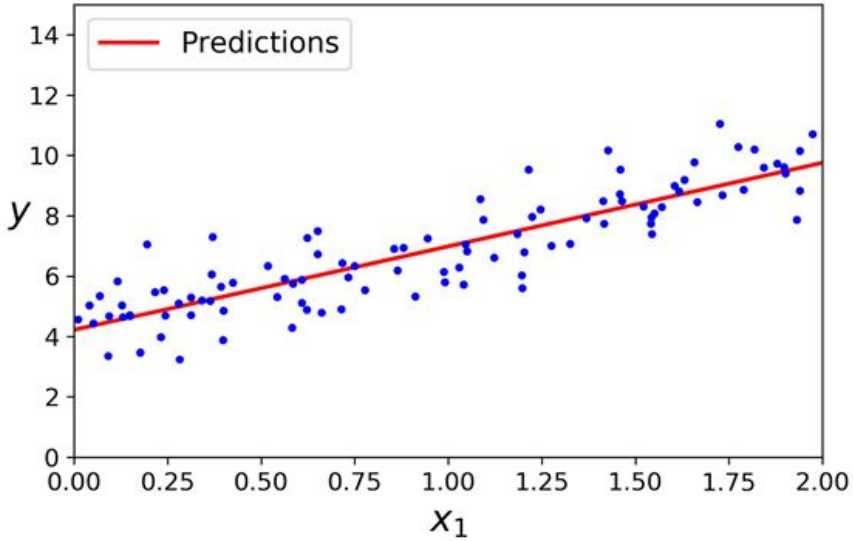
Slide Credit: Amuller, COMS4995

Objective function / دالة الهدف

Mean Squared Error loss

$$J(\theta) = \frac{1}{2n} \sum_{i=1}^{n} (y^{(i)} - \theta^{\top} x^{(i)})^2$$

Linear Regression



$$\hat{y} = w^T \mathbf{x} + b = \sum_{i=1}^{p} w_i x_i + b$$

$$\min_{w \in \mathbb{R}^p, b \in \mathbb{R}} \sum_{i=1}^{n} (w^T \mathbf{x}_i + b - y_i)^2$$

Slide Credit: Amuller, COMS4995

The Normal Equation

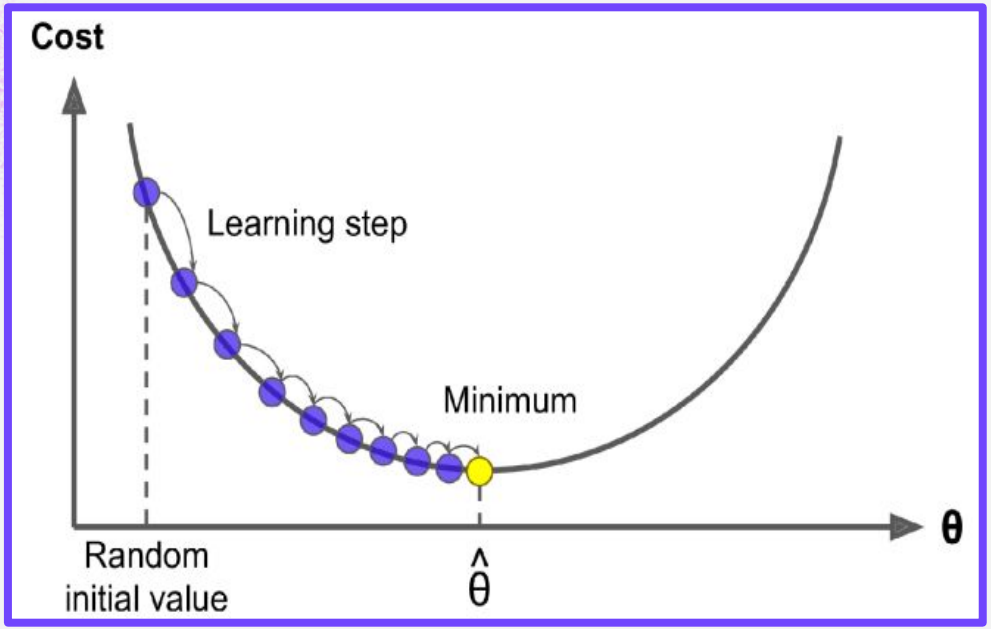
$$\hat{\theta} = (X^T X)^{-1} \quad X^T \quad y$$

“To find the value of θ that minimizes the cost function, there is a closed-form solution”

لكن أين تقع المشكلة ؟

Slide Credit: Aurelien, Hands-on Machine Learning

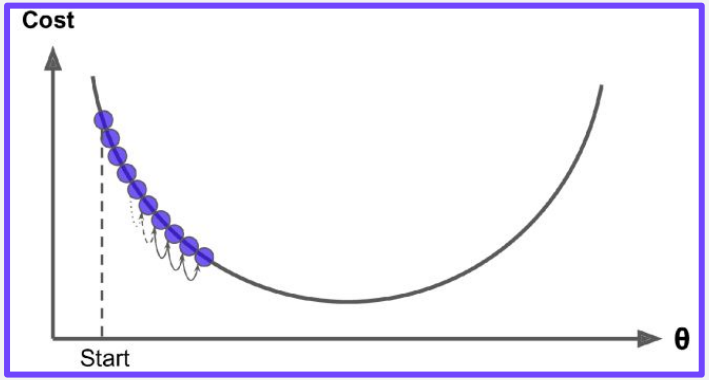
Gradient Descent

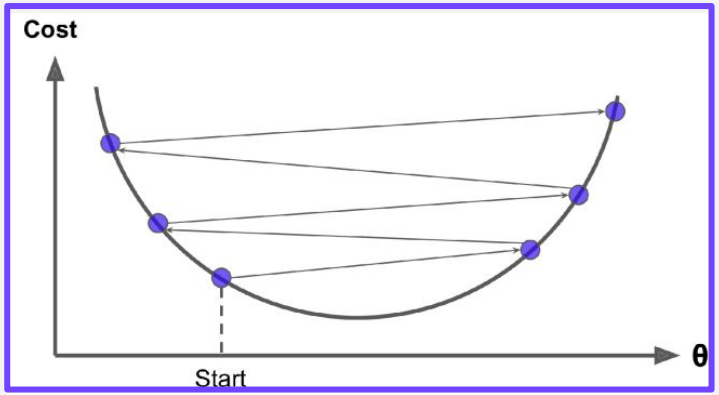


$$\text{repeat until convergence } \{$$
$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$$
$$\}$$

Slide Credit: Aurelien, Hands-on Machine Learning

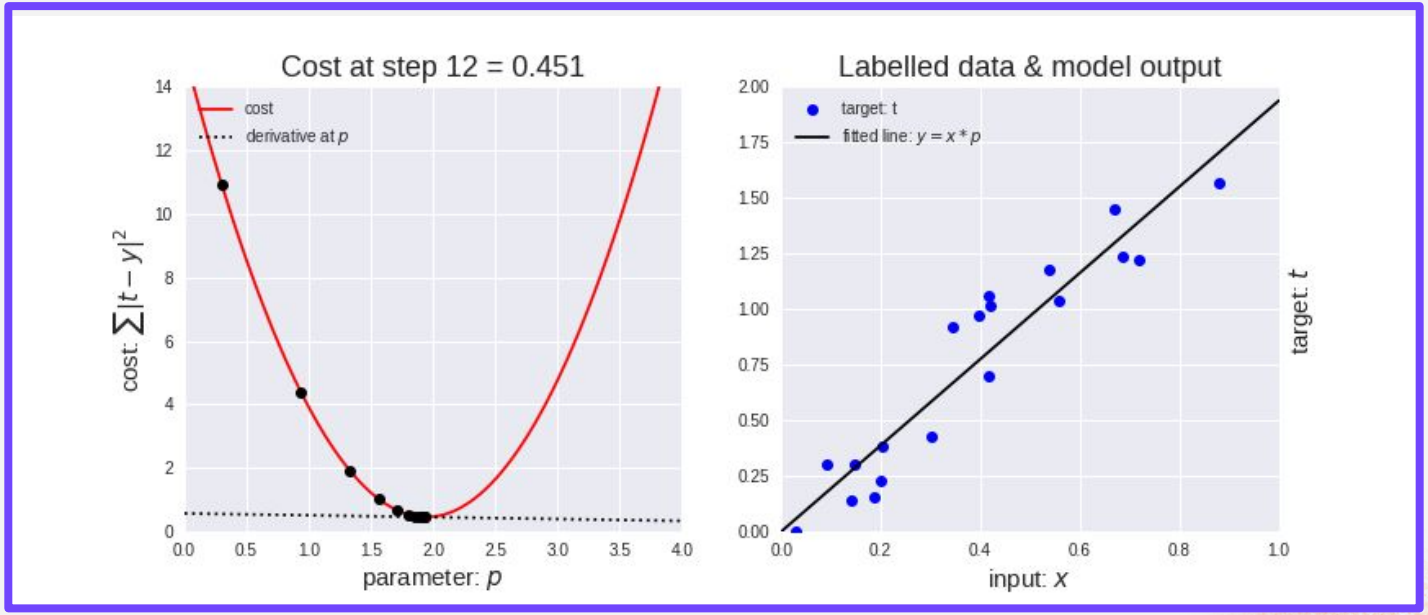
Gradient Descent Learning Rate





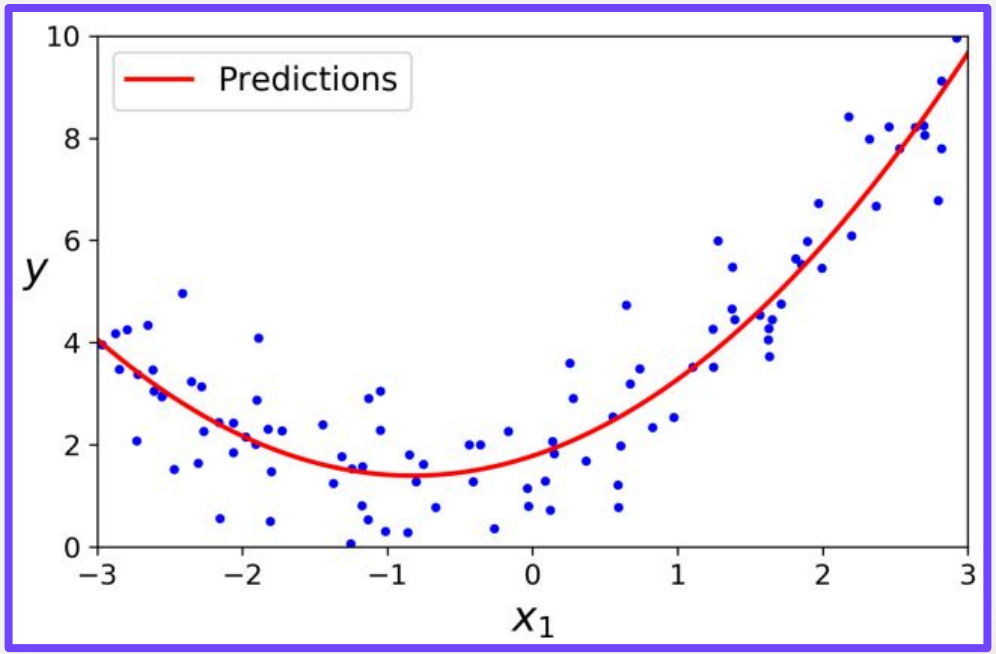
Slide Credit: Aurelien, Hands-on Machine Learning

Gradient Descent



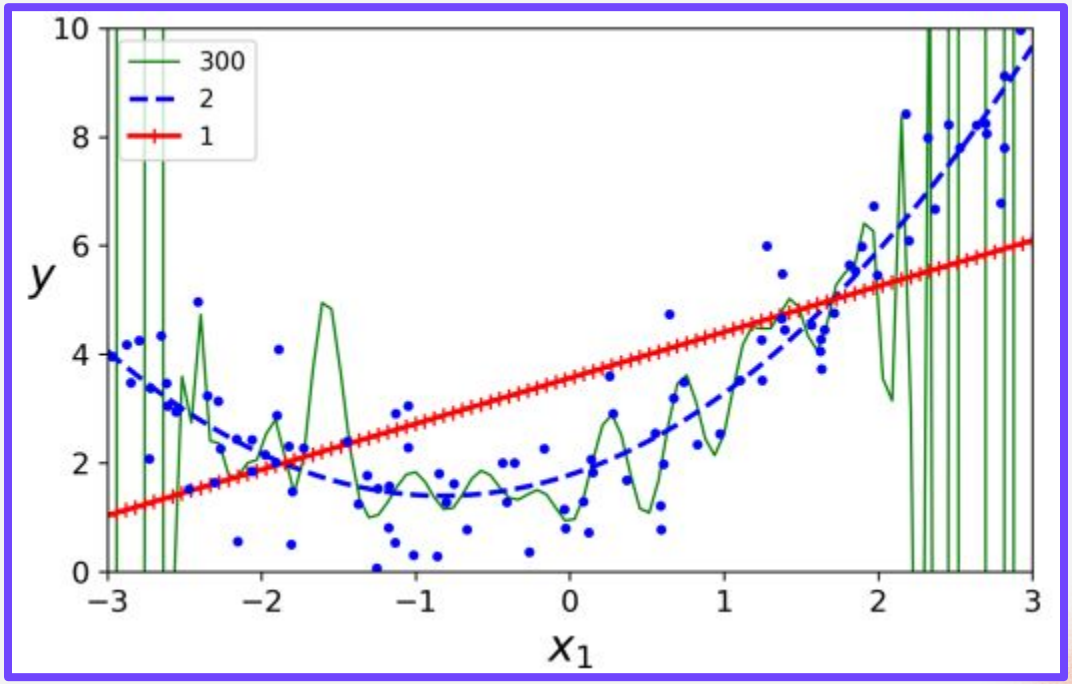
Slide Credit: Github, chengjun

Polynomial Regression



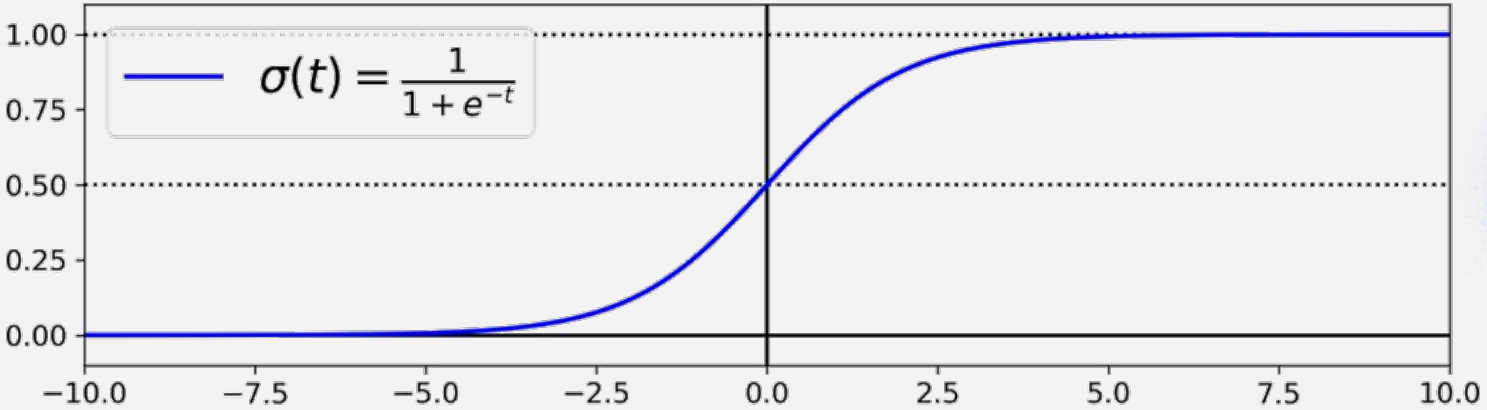
Slide Credit: Aurelien, Hands-on Machine Learning

Polynomial Regression



Slide Credit: Aurelien, Hands-on Machine Learning

Logistic Regression



$$\hat{p} = h_{\boldsymbol{\theta}}(\mathbf{x}) = \sigma(\mathbf{x}^{\intercal}\boldsymbol{\theta})$$

$$\sigma(t) = \frac{1}{1 + \exp(-t)}$$

Slide Credit: Aurelien, Hands-on Machine Learning

Cost Function / دالة التكلفة

$$c(\boldsymbol{\theta}) = \begin{cases} -\log\left(\hat{p}\right) & \text{if } y = 1 \\ -\log\left(1 - \hat{p}\right) & \text{if } y = 0 \end{cases}$$

$$J(\boldsymbol{\theta}) = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}\log\left(\hat{p}^{(i)}\right) + \left(1 - y^{(i)}\right)\log\left(1 - \hat{p}^{(i)}\right)\right]$$

$$\frac{\partial}{\partial\theta_j}\mathrm{J}(\boldsymbol{\theta}) = \frac{1}{m}\sum_{i=1}^{m}\left(\sigma\left(\boldsymbol{\theta}^{\mathsf{T}}\mathbf{x}^{(i)}\right) - y^{(i)}\right)x_j^{(i)}$$

Slide Credit: Aurelien, Hands-on Machine Learning

Softmax Regression

$$s_k(\mathbf{x}) = \mathbf{x}^\top \boldsymbol{\theta}^{(k)}$$

$$\hat{p}_k = \sigma(\mathbf{s}(\mathbf{x}))_k = \frac{\exp\left(s_k(\mathbf{x})\right)}{\sum_{j=1}^{K} \exp\left(s_j(\mathbf{x})\right)}$$

$$\hat{y} = \underset{k}{\operatorname{argmax}}\ \sigma(\mathbf{s}(\mathbf{x}))_k = \underset{k}{\operatorname{argmax}}\ s_k(\mathbf{x}) = \underset{k}{\operatorname{argmax}}\ \left(\left(\boldsymbol{\theta}^{(k)}\right)^\top \mathbf{x}\right)$$

Slide Credit: Aurelien, Hands-on Machine Learning

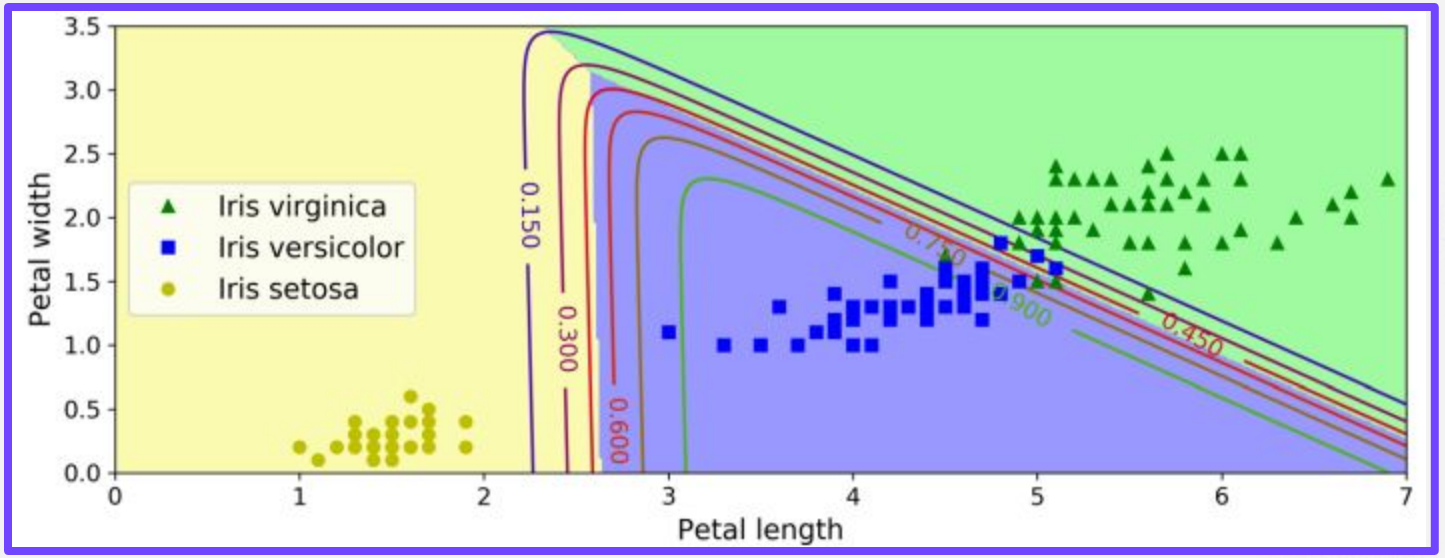
Cost Function / دالة التكلفة

$$J(\Theta) = -\frac{1}{m}\sum_{i=1}^{m}\sum_{k=1}^{K} y_k^{(i)} \log\left(\hat{p}_k^{(i)}\right)$$

$$\nabla_{\theta^{(k)}} J(\Theta) = \frac{1}{m}\sum_{i=1}^{m}\left(\hat{p}_k^{(i)} - y_k^{(i)}\right)\mathbf{x}^{(i)}$$

Slide Credit: Aurelien, Hands-on Machine Learning

Softmax Regression

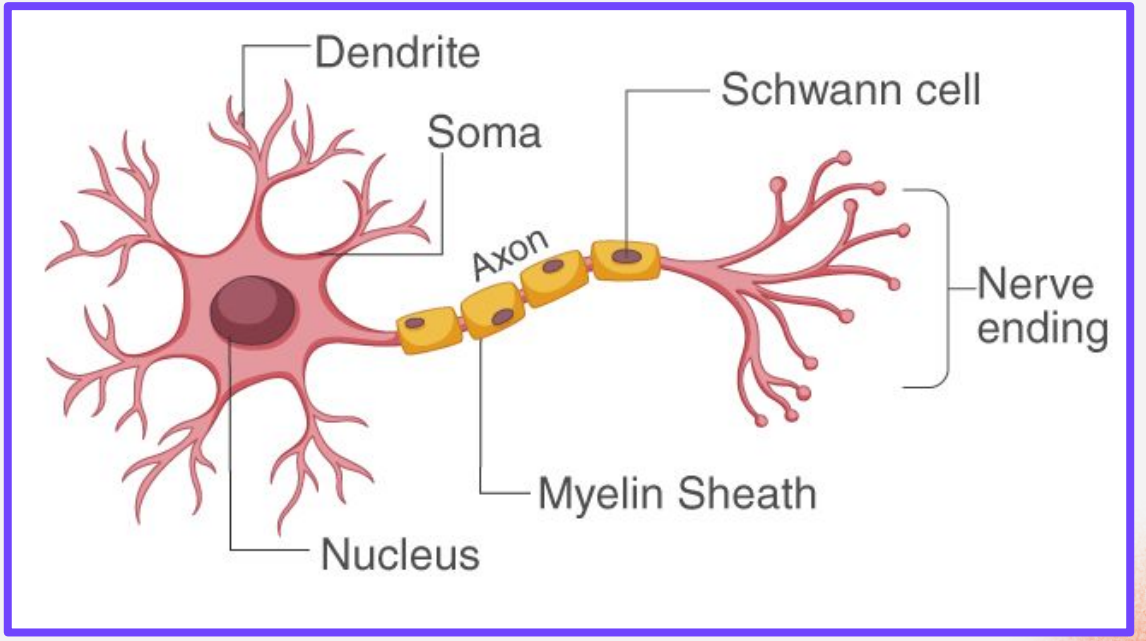


Slide Credit: Aurelien, Hands-on Machine Learning

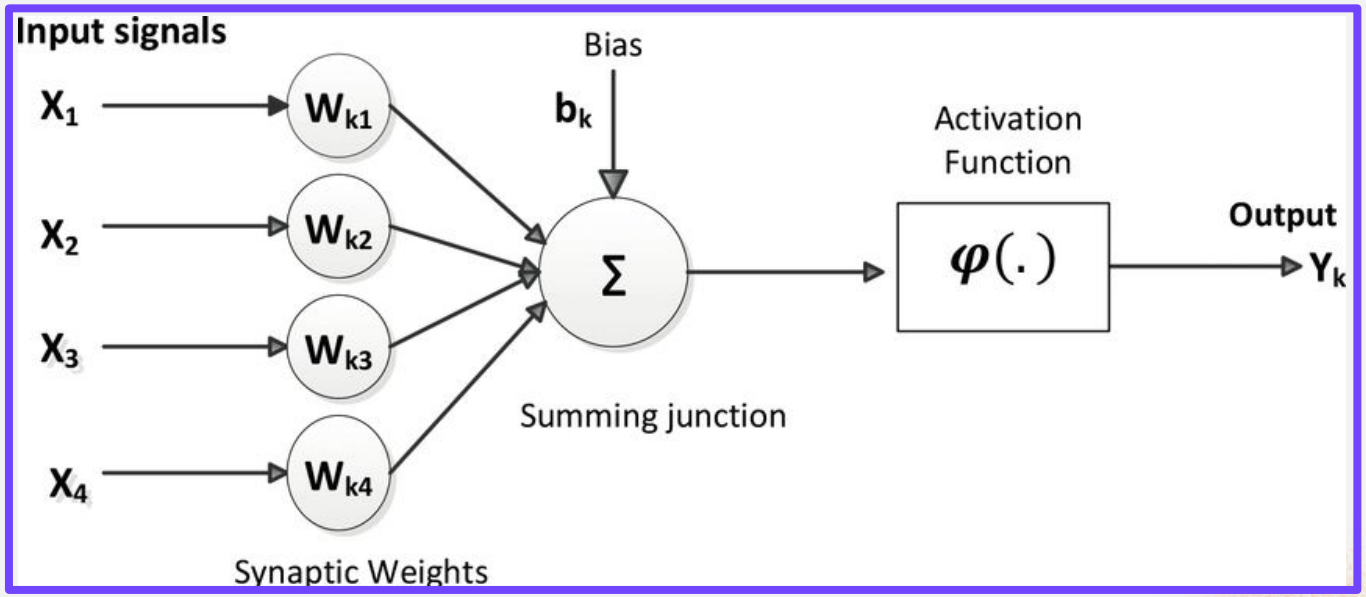
03

أساسيات التعلم العميق
Deep Learning

خلية عصبية نموذجية

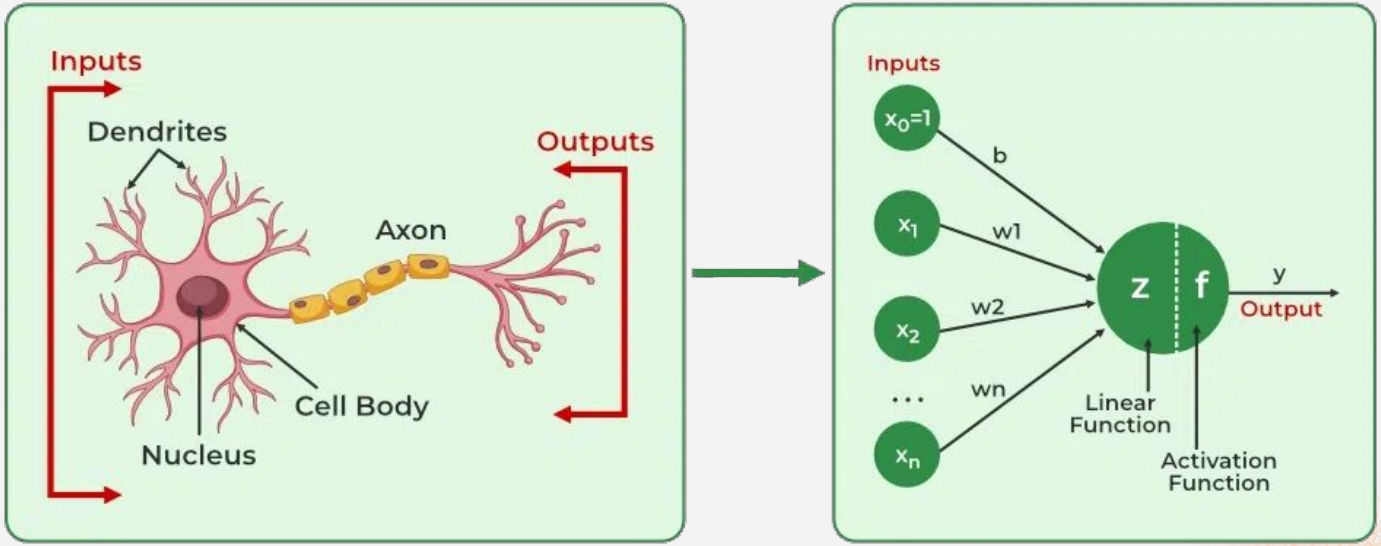


خلية عصبية اصطناعية

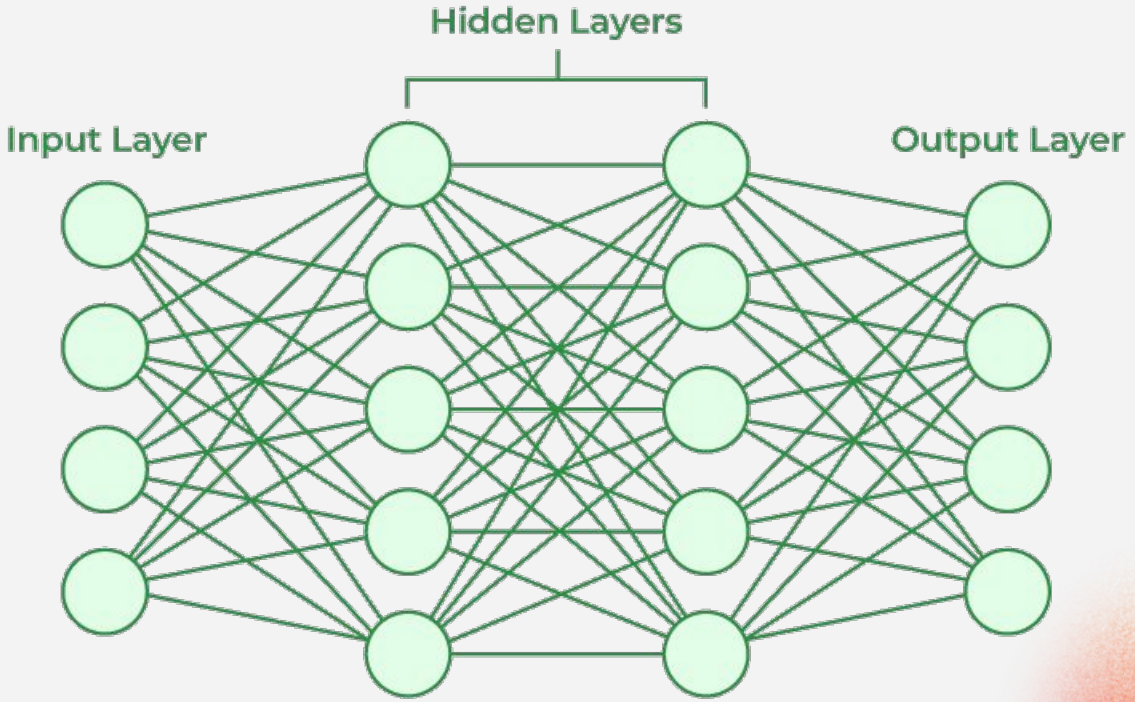


McCulloch Pitts Neuron

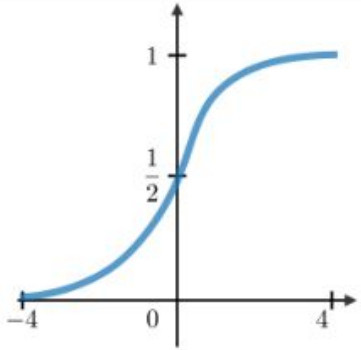
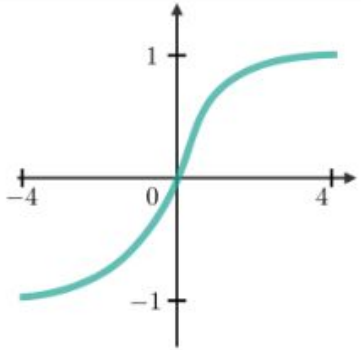
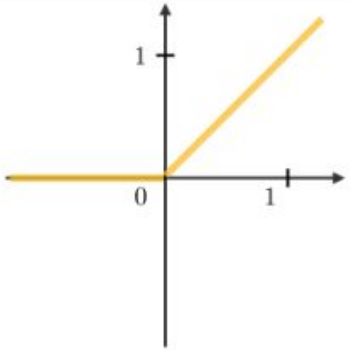
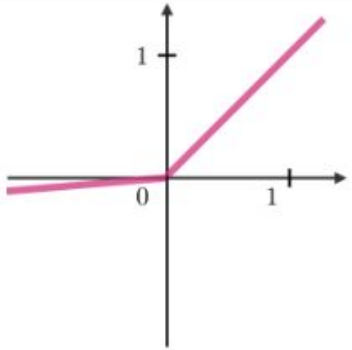
خلية عصبية اصطناعية



Neural Network / الشبكة العصبية

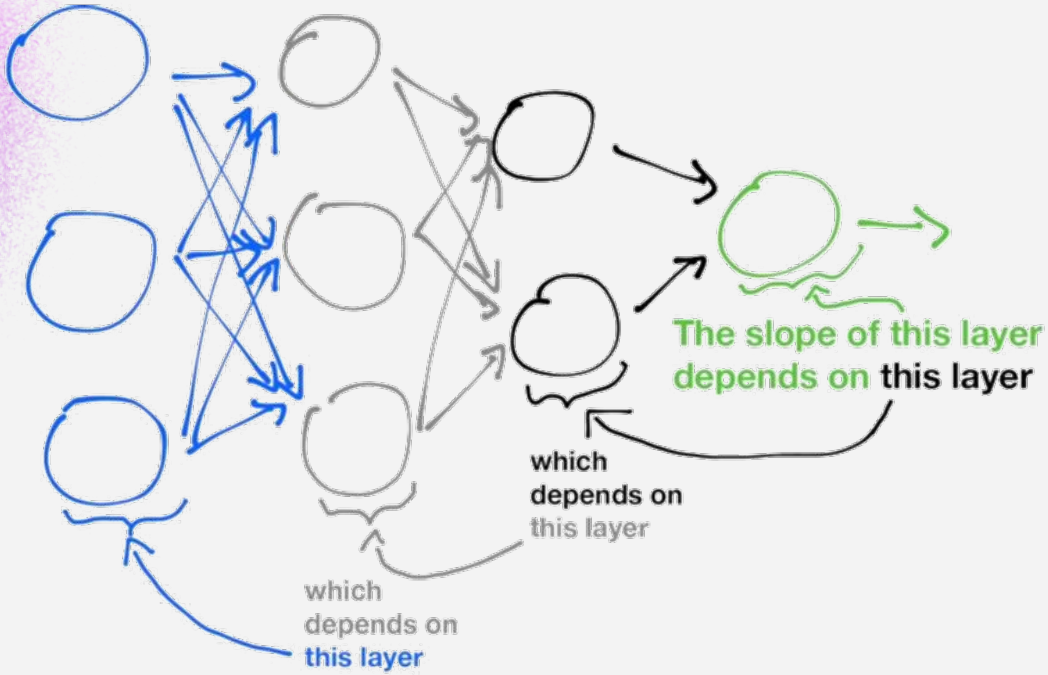


Activation Function

Sigmoid	Tanh	ReLU	Leaky ReLU
$g(z) = \frac{1}{1+e^{-z}}$	$g(z) = \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}}$	$g(z) = \max(0, z)$	$g(z) = \max(\epsilon z, z)$ with $\epsilon \ll 1$
			

Slide Credit: Stanford, Shervine Amidi

Feed Forward/ Backpropagation



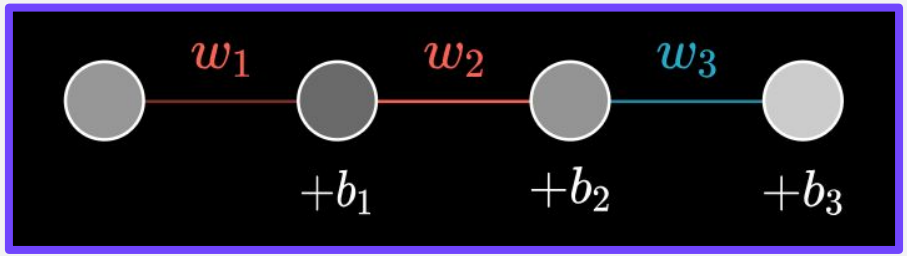
$$\frac{\partial L(z,y)}{\partial w} = \frac{\partial L(z,y)}{\partial a} \times \frac{\partial a}{\partial z} \times \frac{\partial z}{\partial w}$$

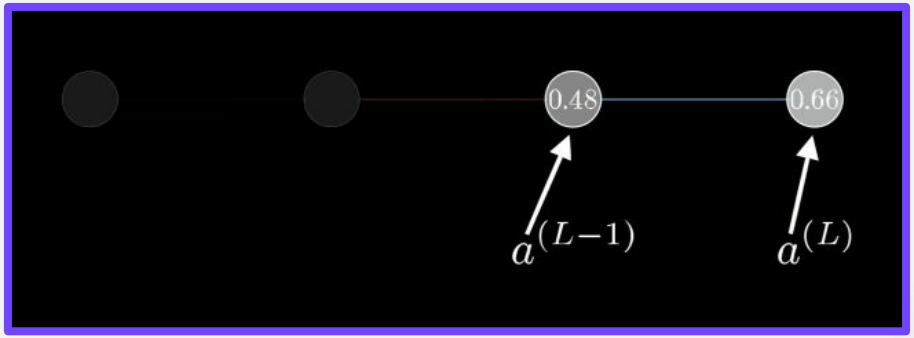
$$w \longleftarrow w - \alpha \frac{\partial L(z,y)}{\partial w}$$

Slide Credit: Siraj Raval

Backpropagation

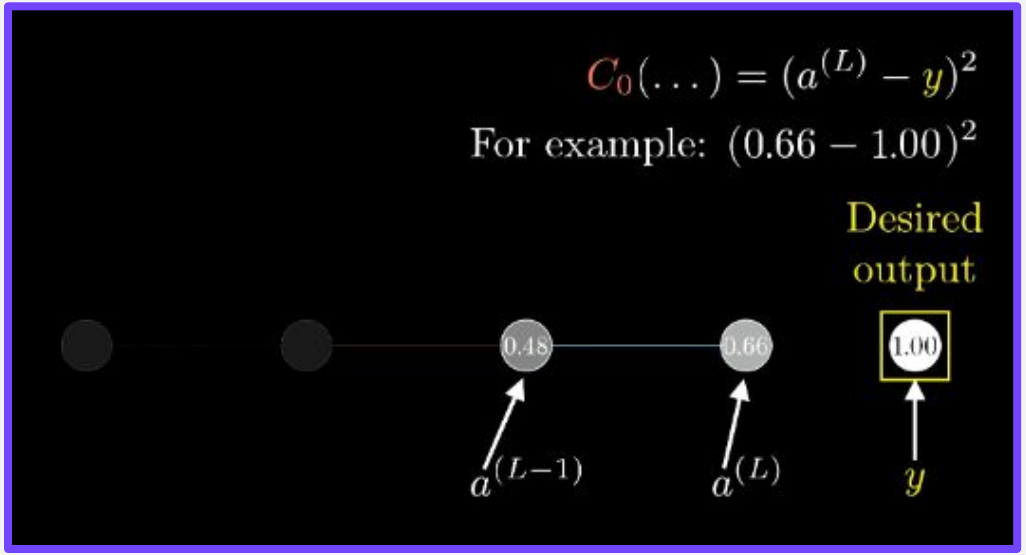
مثال لشبكة بسيطة





Slide Credit: 3Blue1Brown

Backpropagation

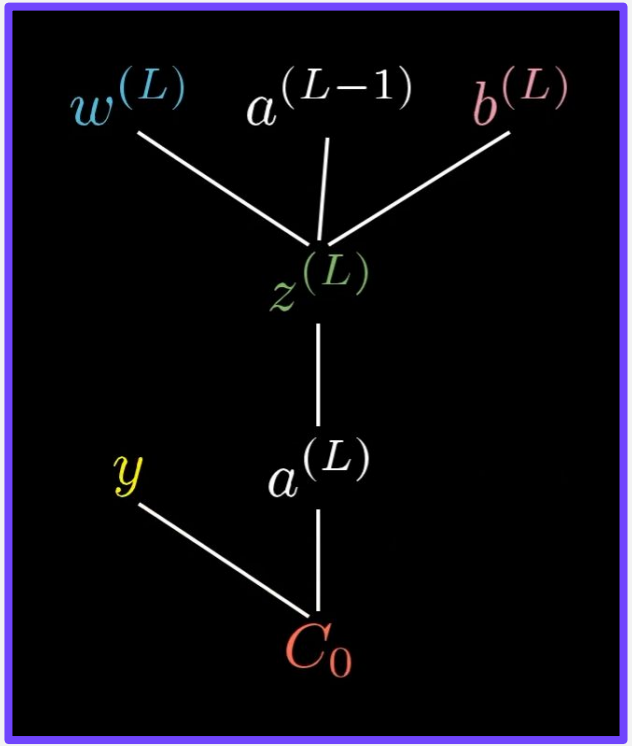


$$z^{(L)} = w^{(L)} a^{(L-1)} + b^{(L)}$$

$$a^{(L)} = \sigma\left(z^{(L)}\right)$$

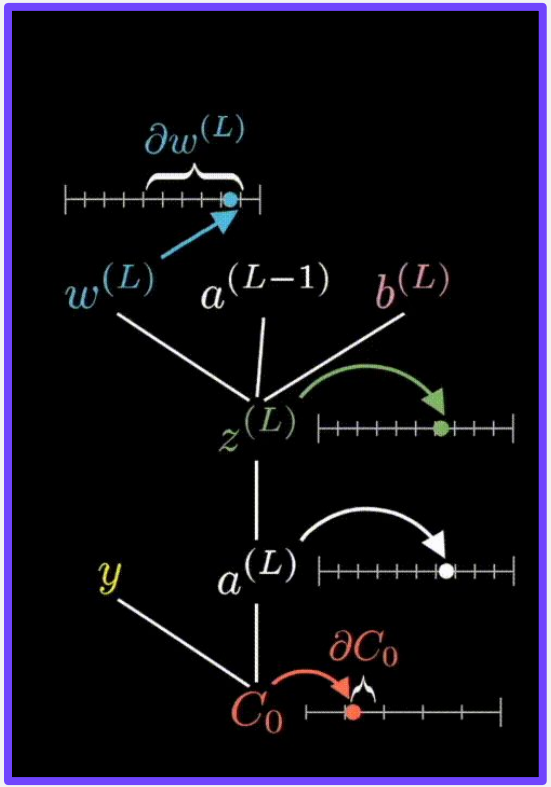
Slide Credit: 3Blue1Brown

Backpropagation



Slide Credit: 3Blue1Brown

Backpropagation



Slide Credit: 3Blue1Brown

Backpropagation: Chain rule

How much does a nudge to $w^{(L)}$ change $z^{(L)}$?

How much does that nudge to $z^{(L)}$ change $a^{(L)}$?

$$\frac{\partial C_0}{\partial w^{(L)}} = \frac{\partial z^{(L)}}{\partial w^{(L)}} \frac{\partial a^{(L)}}{\partial z^{(L)}} \frac{\partial C_0}{\partial a^{(L)}}$$

How much does *that* nudge to $a^{(L)}$ change C_0?

Slide Credit: 3Blue1Brown

Backpropagation: Chain rule

$$z^{(L)} = w^{(L)} a^{(L-1)} + b^{(L)} \quad \longrightarrow \quad \frac{\partial z^{(L)}}{\partial w^{(L)}} = a^{(L-1)}$$

$$a^{(L)} = \sigma\left(z^{(L)}\right) \quad \longrightarrow \quad \frac{\partial a^{(L)}}{\partial z^{(L)}} = \sigma'\left(z^{(L)}\right)$$

$$C_0 = \left(a^{(L)} - y\right)^2 \quad \longrightarrow \quad \frac{\partial C_0}{\partial a^{(L)}} = 2\left(a^{(L)} - y\right)$$

Slide Credit: 3Blue1Brown

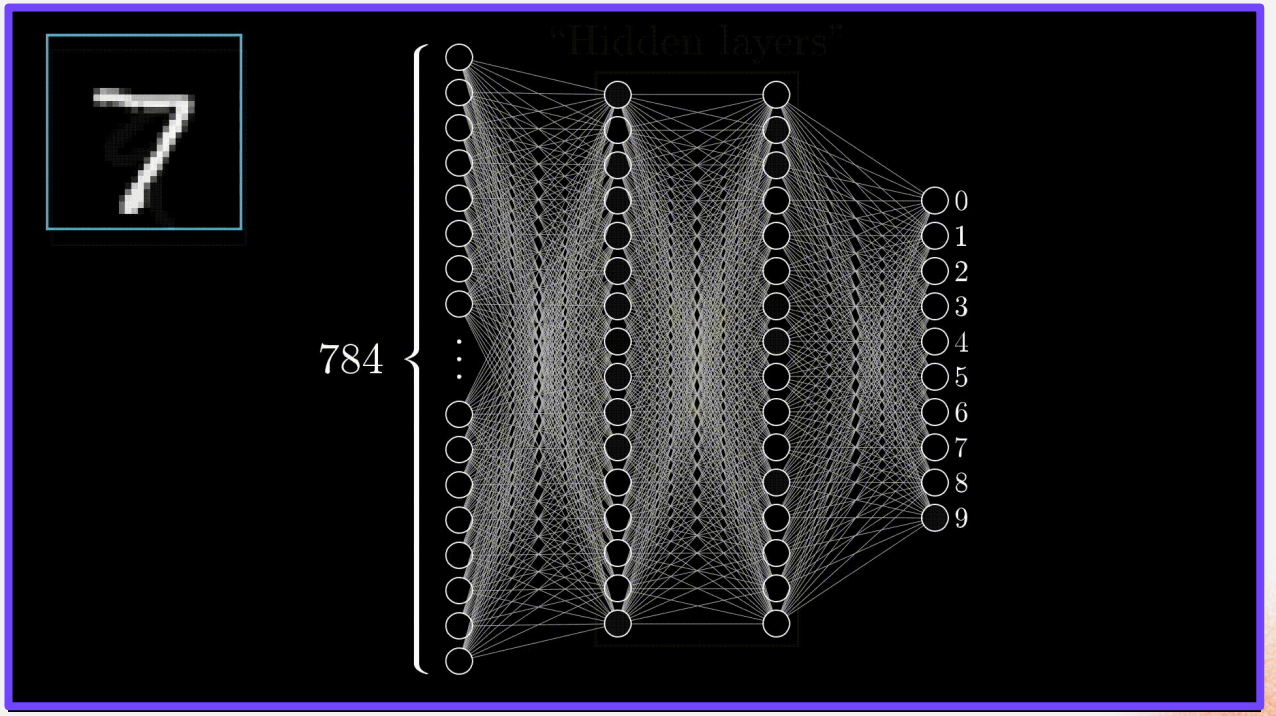
Backpropagation: Chain rule

$$\frac{\partial C_0}{\partial w^{(L)}} = a^{(L-1)} \sigma'\left(z^{(L)}\right) 2\left(a^{(L)} - y\right)$$

Use Gradient Descent to update the weights

Slide Credit: 3Blue1Brown

مثال لتصنيف الأرقام



Slide Credit: 3Blue1Brown

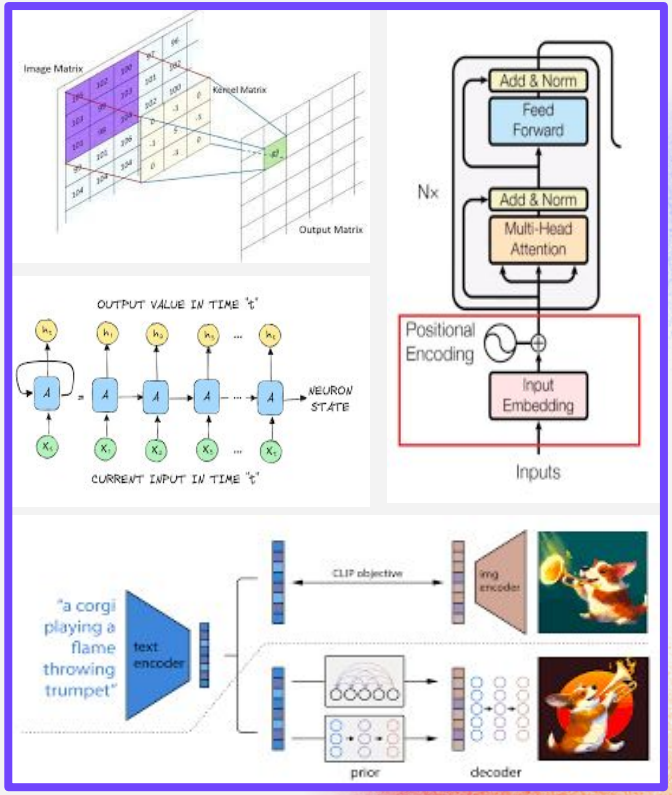
DL Frameworks

TensorFlow

PyTorch

الخطوات التالية

- Convolutional Neural Networks (CNN)
- Recurrent Neural Networks (RNN)
- Diffusion Models
- Transformers (NLP)
- Reinforcement Learning



مصادر


O'REILLY
Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow
Concepts, Tools, and Techniques to Build Intelligent Systems
Aurélien Géron
Andriy Burkov
THE HUNDRED-PAGE MACHINE LEARNING BOOK
DEEP LEARNING
Ian Goodfellow, Yoshua Bengio, and Aaron Courville
DEEP LEARNING with Python
François Chollet


Machine Learning Specialization
Deep Learning Specialization


Google Scholar
NON SOLUS
ELSEVIER
arXiv

شكرا على حسن استماعكم