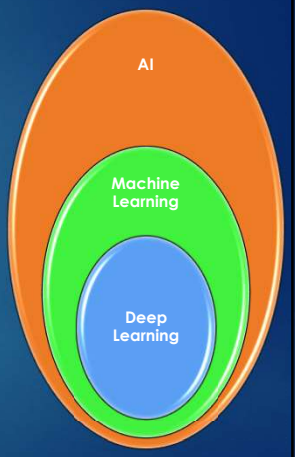
Introduction to deep learning

FAKHRLEIDN SAEED

1

What Deep Learning Means

- **Artificial intelligence** is the ability of a computer to perform tasks commonly associated with intelligent beings.
- **Machine learning** is the study of algorithms that learn from examples and experience instead of relying on hard-coded rules and make predictions on new data.
- **Deep learning** is a subfield of machine learning focusing on learning data representations as successive layers of increasingly meaningful representations.



2

The differences between ML and DL (1/4)

- **A completely different paradigm:**
- **Flexible architectures:**
- **Autonomous feature definition:**

3

The differences between ML and DL (2/4)

- **A completely different paradigm:**
  - ML is a **set of many** different techniques that enable a computer to learn from data and to use what it learns to provide an answer, often in the form of a prediction.
  - ML relies on **different paradigms** such as using statistical analysis, finding analogies in data, using logic, and working with symbols.
  - DL used a single technique, which mimics human brain functionality (neural network). It processes data using computing units, called **neurons**, arranged into ordered sections, called **layers**.

4

The differences between ML and DL (3/4)

- **Flexible architectures:**
  - ML solutions offer many knobs (adjustments) called ***hyperparameters*** that you tune to optimize algorithm learning from data.
  - DL solutions use ***hyperparameters***, too, but they also use multiple user-configured ***layers*** (the user specifies number and type).
  - In fact, depending on the resulting neural network, the number of layers can be quite large and form unique neural networks capable of specialized learning: Some can learn to recognize images, while others can detect and parse voice commands.
  - The point is that the term deep is appropriate; it refers to the large number of layers potentially used for analysis. The architecture consists of the ensemble of different neurons and their arrangement in layers in a deep learning solution.

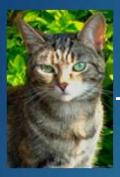
5

The differences between ML and DL (4/4)

- **Autonomous feature definition:**
  - ML solutions require human intervention to succeed.
    - To process data correctly, analysts and scientist use a lot of their own knowledge to develop working algorithms.
    - Creating the right information for a machine learning algorithm is called ***feature creation***, which is a time-consuming activity.
  - Deep learning doesn't require humans to perform any feature-creation activity because, thanks to its many layers, it defines its own best features.
  - That's also why deep learning outperforms machine learning in otherwise very difficult tasks such as recognizing voice and images, understanding text, or beating a human champion at the Go game.

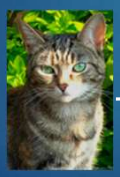
6

"Traditional" machine learning:



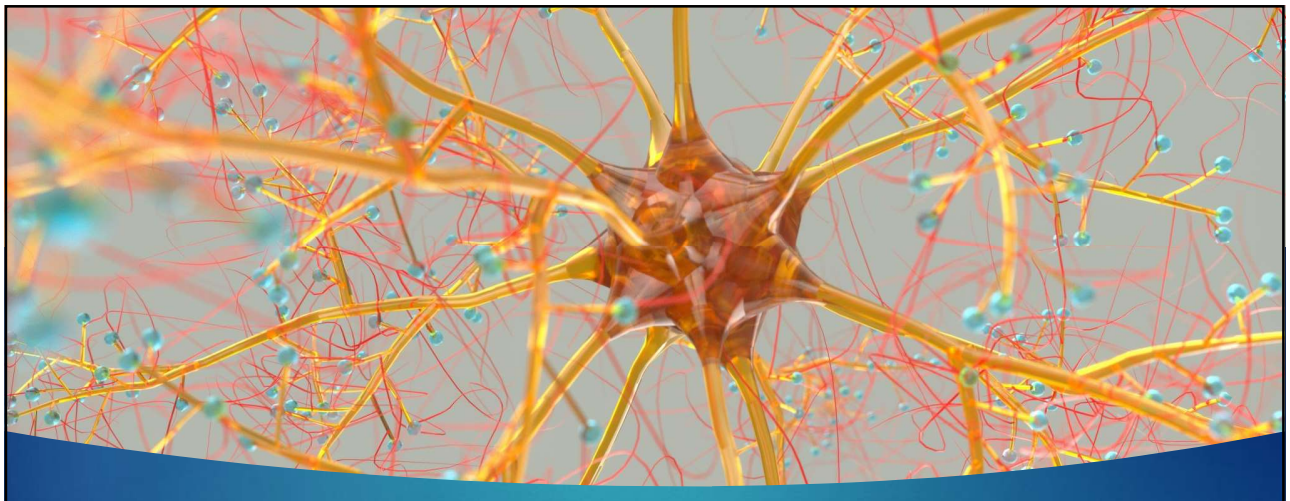
handcrafted features → learned classifier → cat

Deep, "end-to-end" learning:



learned low-level features → learned mid-level features → learned high-level features → learned classifier → cat

7



Neural Network

8

Defining Artificial Neural Networks

- A artificial neural network is structured like the human brain and consists of artificial neurons, also known as nodes. These nodes are stacked next to each other in three layers:
  - The input layer
  - The hidden layer(s)
  - The output layer

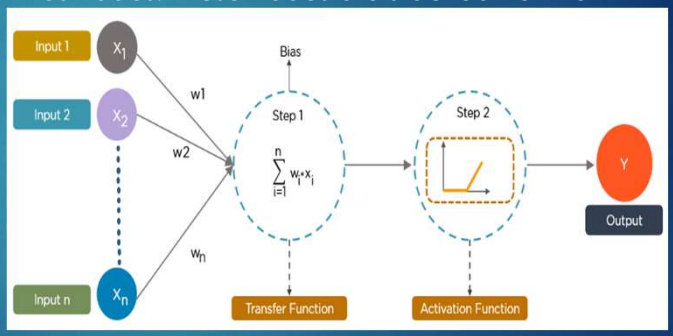
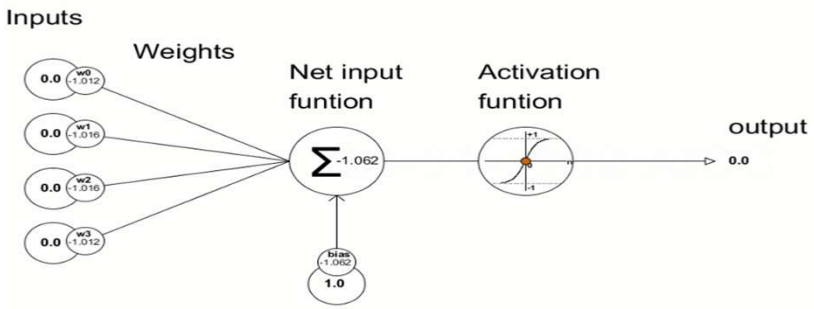


IMAGE :https://www.simplilearn.com/tutorials/deep-learning-tutorial/deep-learning-algorithm

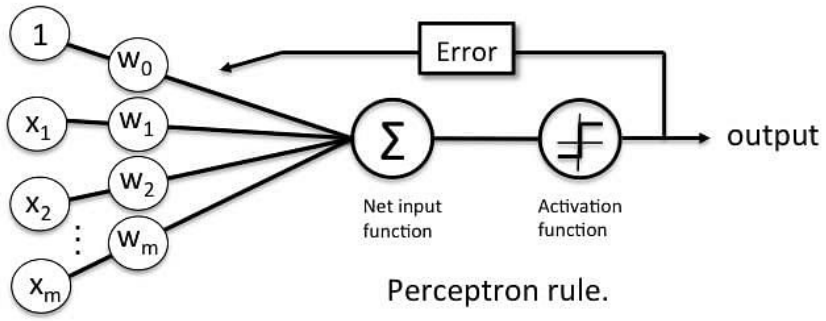
9



Example

https://www.mql5.com/en/blogs/post/724245

10



Perceptron

https://www.simplilearn.com/tutorials/deep-learning-tutorial/perceptron

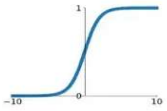
11

Structure of ANN

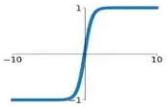
- A better way to understand ANNs is to think of their structure as a combination of the hyper parameters in the following list:
  - The number of hidden layers
  - The number of neurons in each hidden layer
  - The initial weights of edges connecting pairs of neurons
  - The activation function (to add non-linearity into a neural network.)
  - A cost (loss) function
  - An optimizer (used with the cost function)
  - The learning rate (a small number)
  - The dropout rate (optional)

12

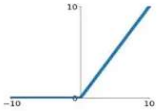
Sigmoid
$\sigma(x) = \frac{1}{1+e^{-x}}$



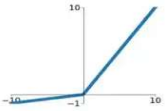
tanh
$\tanh(x)$



ReLU
$\max(0, x)$



Leaky ReLU
$\max(0.1x, x)$



Maxout
$\max(w_1^T x + b_1, w_2^T x + b_2)$

ELU
$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$

Activation functions

https://medium.com/@shrutijadon/survey-on-activation-functions-for-deep-learning-9689331ba092

13

Cost (loss) function

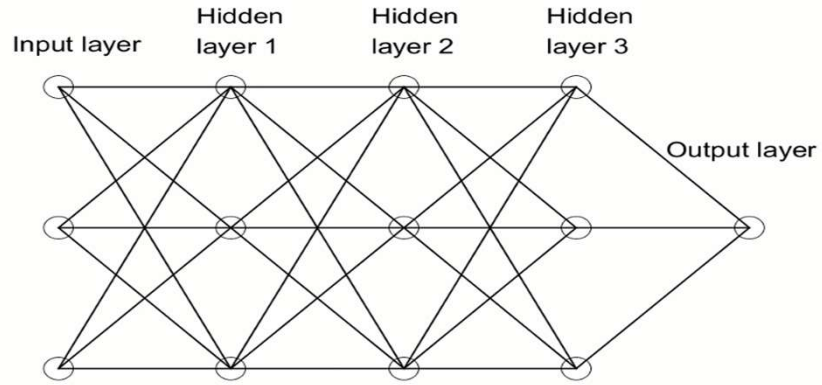
- **Cost Function** is used to measure just how wrong the model is in finding a relation between the input and output. It tells you how badly your model is behaving/predicting
- **Gradient Descent** is an algorithm that is used to optimize the cost function or the error of the model. It is used to find the minimum value of error possible in your model.

Cost Function (J) = $\frac{1}{n}\sum_{i=0}^{n}(y^i - (mx^i + b))^2$

Gradient Descent $(\frac{\partial J}{\partial \theta})$ = $\begin{bmatrix} \frac{1}{N}\sum_{i=0}^{n}(-2x_i(y_i - (mx_i + b))) \\ \frac{1}{N}\sum_{i=0}^{n}(-2(y_i - (mx_i + b))) \end{bmatrix}$

https://www.simplilearn.com/tutorials/machine-learning-tutorial/cost-function-in-machine-learning

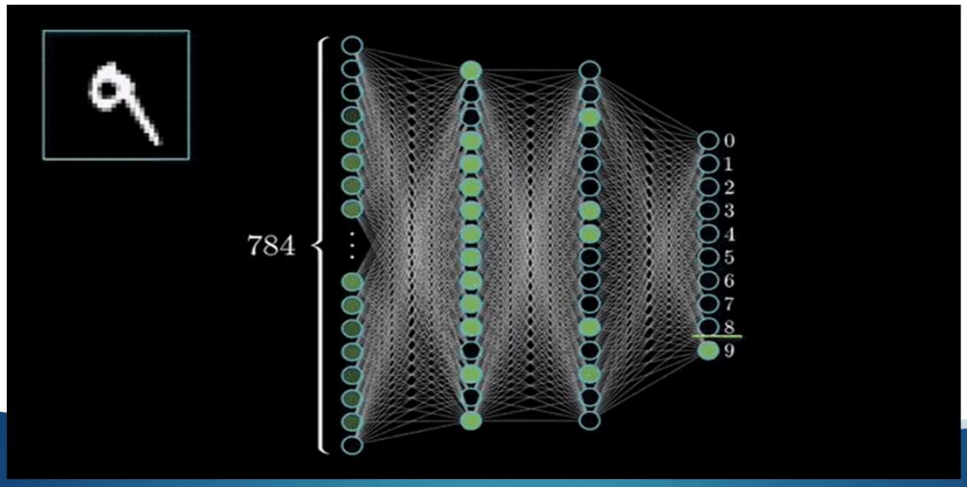
14



Deep Neural Network

https://www.mql5.com/en/blogs/post/724245

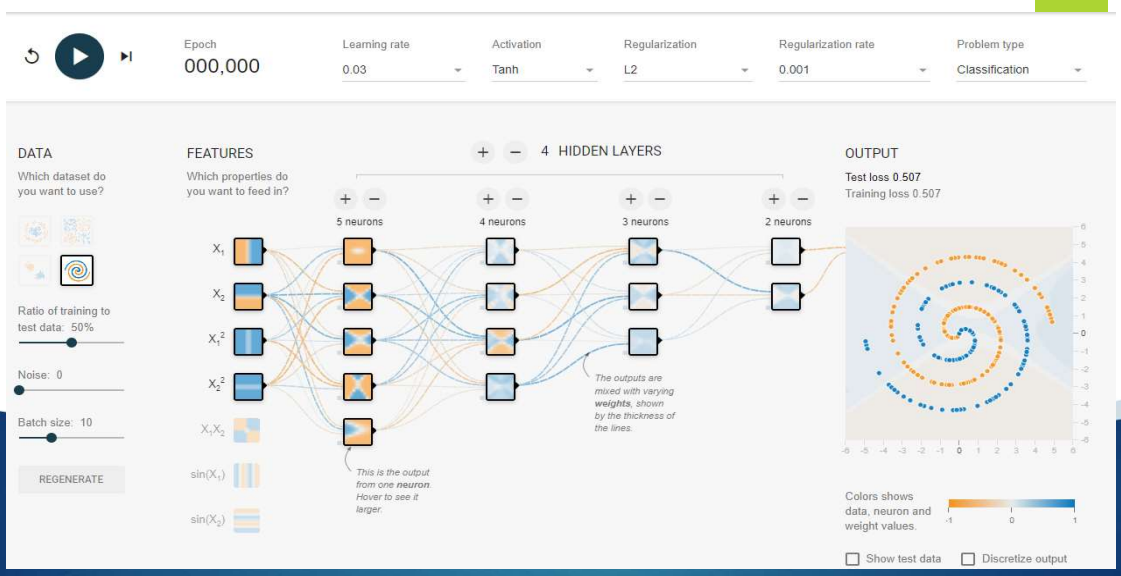
15



Example

https://gfycat.com/gifs

16



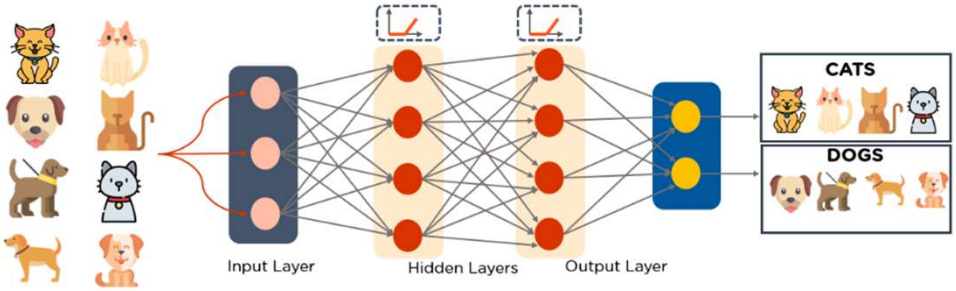
https://playground.tensorflow.org/

17

Deep learning architectures

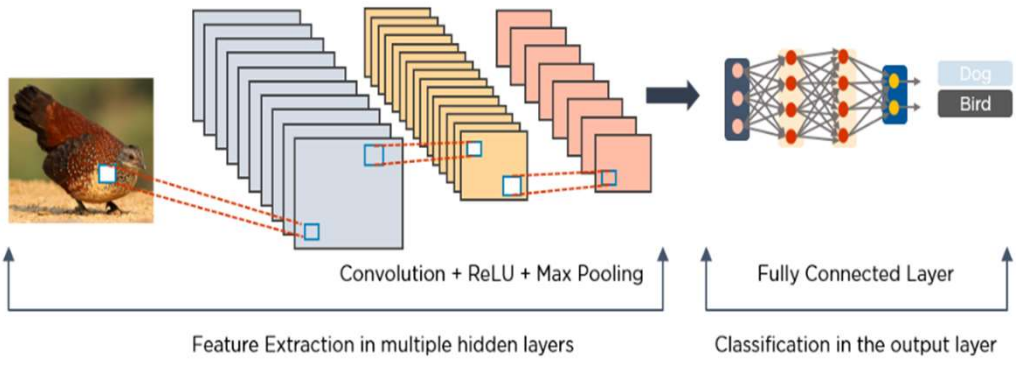
- Multilayer Perceptrons (**MLPs**)
- Convolutional Neural Networks (**CNNs**)
- Recurrent Neural Networks (**RNNs**)
- Long Short Term Memory Networks (**LSTMs**)
- Generative Adversarial Networks (**GANs**)
- Deep reinforcement learning is a combination of reinforcement learning with these architectures.

18



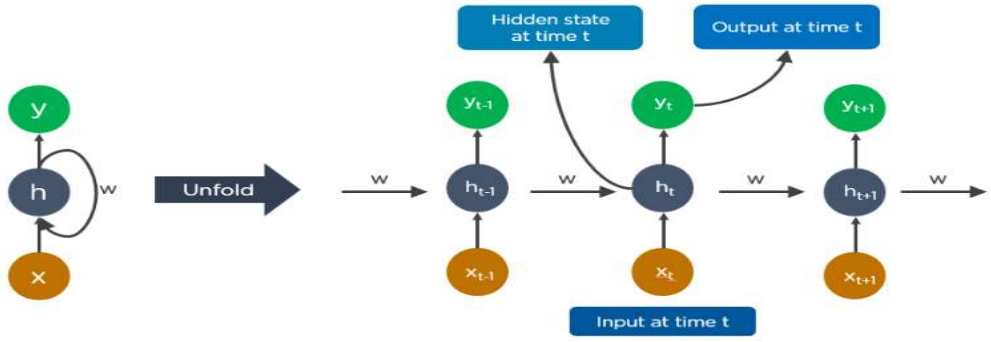
Multilayer Perceptrons (MLPs)

19



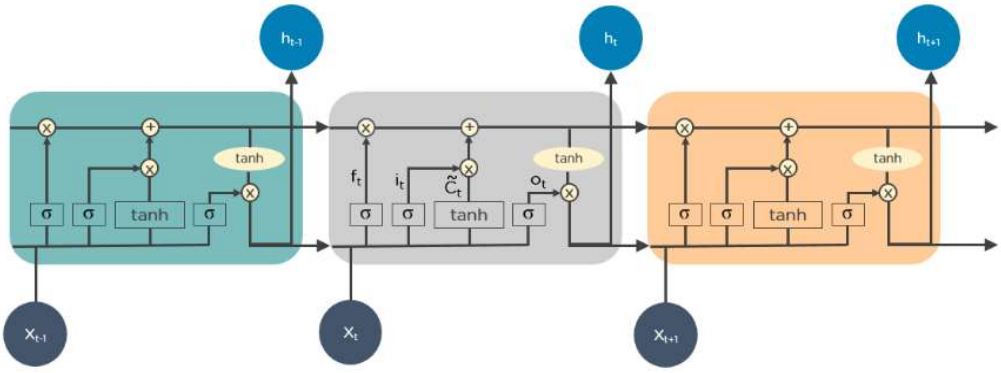
Convolutional Neural Networks (CNNs)

20



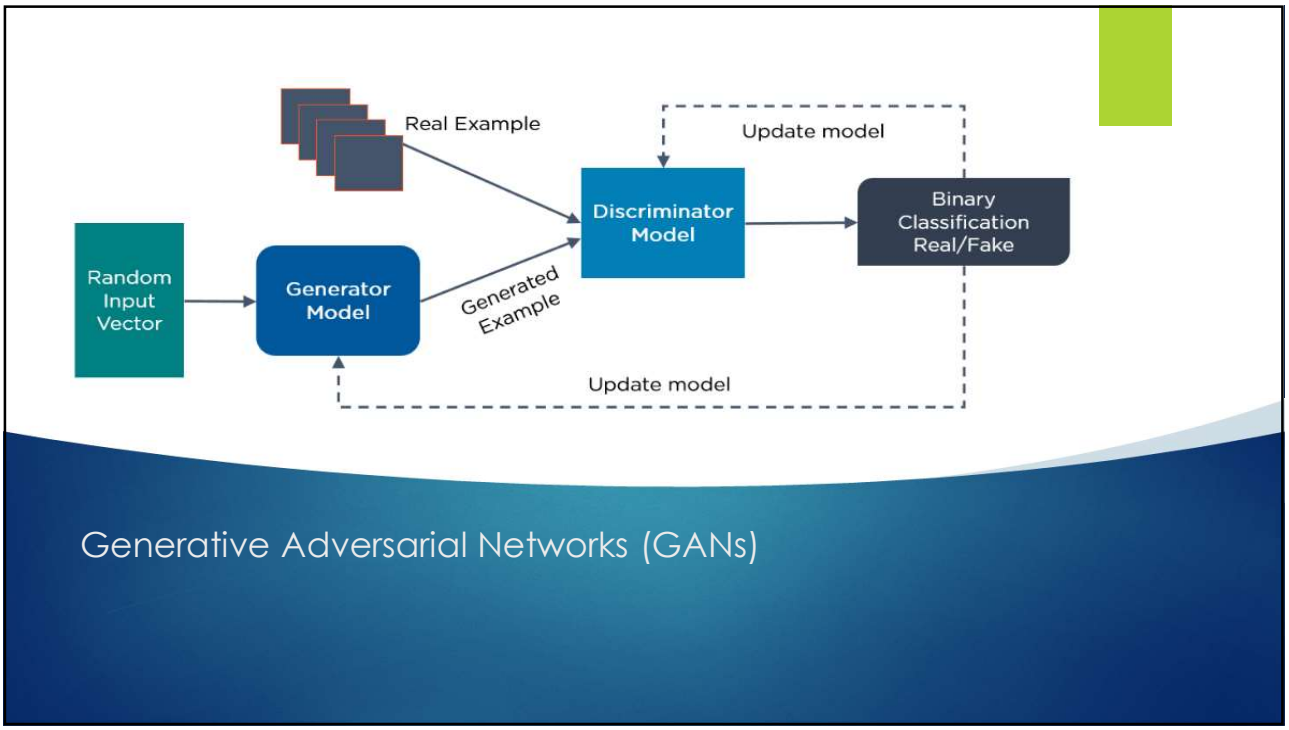
Recurrent Neural Networks (RNNs)

21



Long Short Term Memory Networks (LSTMs)

22



Generative Adversarial Networks (GANs)

23

Examples of Deep Learning Applications

Self Driving Cars	News Aggregation and Fraud News Detection	Natural Language Processing	Virtual Assistants	Entertainment
Visual Recognition	Fraud Detection	Healthcare	Personalisations	Detecting Developmental Delay in Children
Colourisation of Black and White images	Adding sounds to silent movies	Automatic Machine Translation	Automatic Handwriting Generation	Automatic Game Playing
Language Translations	Pixel Restoration	Photo Descriptions	Demographic and Election Predictions	Deep Dreaming

24

Deep Learning frameworks

- Deep learning frameworks are software libraries that provide tools and libraries for training and deploying deep learning models. Some popular deep learning frameworks include:
  - **TensorFlow**: Developed by Google, is an open-source framework for building and training machine learning models. It is widely used for a variety of tasks, including image and language processing.
  - **PyTorch**: Developed by Facebook, is an open-source framework that is built for flexibility and speed. It is popular for its ability to create dynamic computational graphs and its support for distributed training.
  - **Keras**: a high-level deep learning framework that is built on top of other frameworks, such as TensorFlow, Theano, and CNTK. It is designed to be easy to use and allows users to build complex deep learning models with minimal code.
  - **MXNet**: an open-source deep learning framework developed by Amazon. It is designed for flexibility and efficient use of resources, and it supports a wide range of hardware platforms and programming languages.
  - **Caffe**: an open-source deep learning framework developed by the Berkeley Vision and Learning Centre (BVLC). It is widely used for image classification tasks and has a large user community.

25

Popular Deep Learning Services in the Cloud

- Infrastructure as a service (IaaS) model:
  - Google TPU, AWS GPU and Azure GPU
- Platform as a service (PaaS) : provide the hardware, as well as software services for managing deep learning pipelines, from data ingestion to production deployment and real-world inference.
  - SageMaker from AWS, Cloud AI from Google, Azure Machine Learning
- Deep learning frameworks Support:
  - TensorFlow, PyTorch, Keras, MXNet and Deep Java Library

26

Pre-Trained AI Services

- Most cloud platforms provide pre-trained, pre-optimized AI services for many applications including:
  - Image classification
  - Object recognition
  - Video data extraction
  - Language translation
  - Speech synthesis
  - Recommendation engines

27

Transfer learning

- **Transfer learning** is a machine learning technique that involves reusing a pre-trained model on a new task. It is based on the idea that the knowledge learned by a model during training on one task can be useful for learning other related tasks.
- The main **advantage** of transfer learning is that it can significantly **reduce** the amount of labeled data and computational resources needed to train a good model on a new task. This is because the pre-trained model has already learned useful features and representations from the original task, and these can be transferred and fine-tuned for the new task.
- There are several different approaches to transfer learning, including using
  - a pre-trained model as a fixed feature extractor.
  - fine-tuning the model on the new task.
  - using the model as a initialization for training a new model on the new task.
- The specific approach used will depend on the nature of the new task and the availability of labeled data for the new task.

28