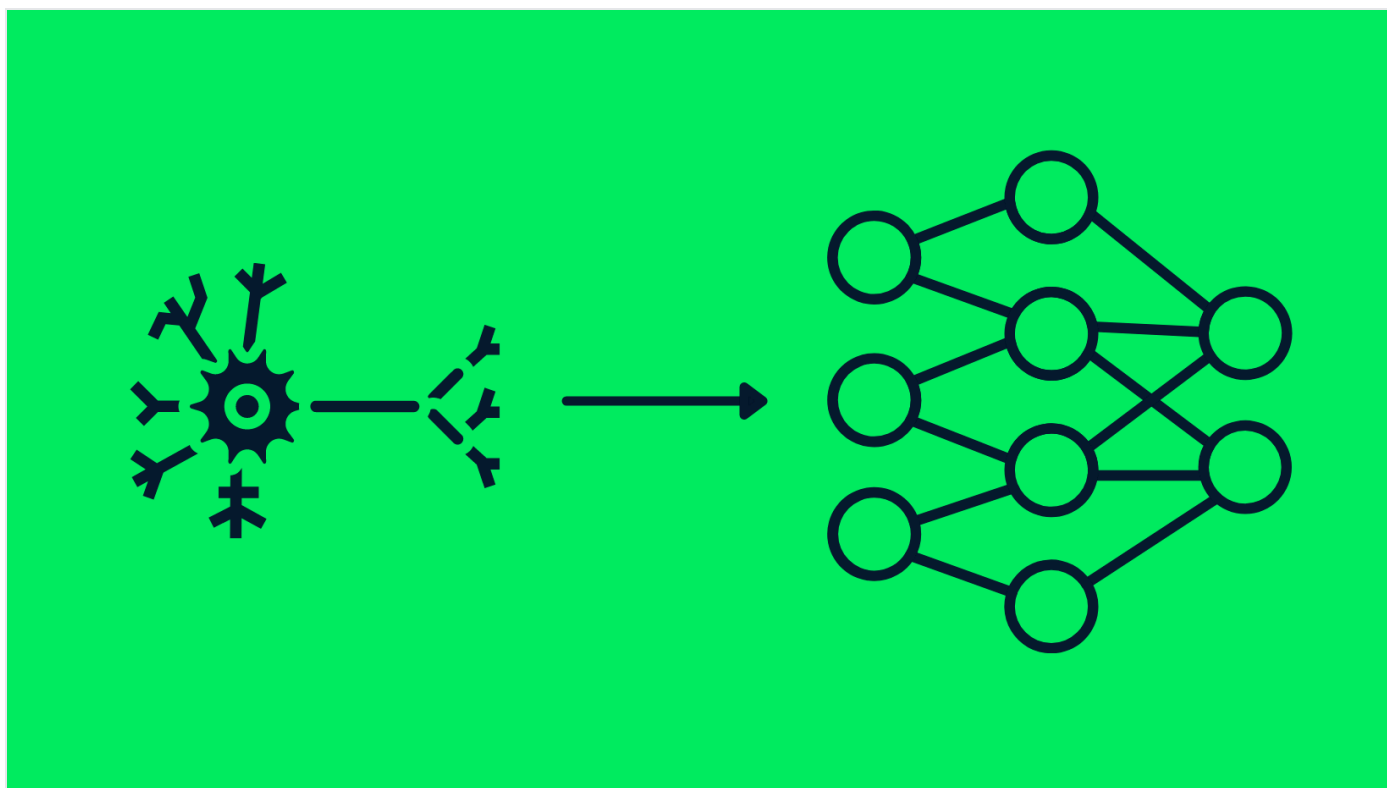


Deep learning is a type of machine learning that teaches computers to perform tasks by learning from examples, much like humans do. Imagine teaching a computer to recognize cats: instead of telling it to look for whiskers, ears, and a tail, you show it thousands of pictures of cats. The computer finds the common patterns all by itself and learns how to identify a cat. This is the essence of deep learning.

In technical terms, deep learning uses something called "neural networks," which are inspired by the human brain. These networks consist of layers of interconnected nodes that process information. The more layers, the "deeper" the network, allowing it to learn more complex features and perform more sophisticated tasks.



The Similarity Between Neurons and Neural Networks

The Evolution of Machine Learning to Deep Learning

What is machine learning?

Machine learning is itself a subset of **artificial intelligence** (AI) that enables computers to learn from data and make decisions without explicit programming. It encompasses various techniques and algorithms that allow systems to recognize patterns, make predictions, and improve performance over time. You can explore the **difference between machine learning and AI** in a separate article.

How deep learning differs from traditional machine learning

While machine learning has been a transformative technology in its own right, deep learning takes it a step further by automating many of the tasks that typically require human expertise.

Deep learning is essentially a specialized subset of machine learning, distinguished by its use of neural networks with three or more layers. These neural networks attempt to simulate the behavior of the human brain—albeit far from matching its ability—in order to "learn" from large amounts of data. You can explore [machine learning vs deep learning](#) in more detail in a separate post.

The importance of feature engineering

Feature engineering is the process of selecting, transforming, or creating the most relevant variables, known as "features," from raw data to use in machine learning models.

For example, if you're building a weather prediction model, the raw data might include temperature, humidity, wind speed, and barometric pressure. Feature engineering would involve determining which of these variables are most important for predicting the weather and possibly transforming them (e.g., converting temperature from Fahrenheit to Celsius) to make them more useful for the model.

In traditional machine learning, feature engineering is often a manual and time-consuming process that requires domain expertise. However, one of the advantages of deep learning is that it can automatically learn relevant features from the raw data, reducing the need for manual intervention.

Why is Deep Learning Important?

The reasons why deep learning has become the industry standard:

- **Handling unstructured data:** Models trained on structured data can easily learn from unstructured data, which reduces time and resources in standardizing data sets.
- **Handling large data:** Due to the introduction of graphics processing units (GPUs), deep learning models can process large amounts of data with lightning speed.
- **High Accuracy:** Deep learning models provide the most accurate results in computer visions, natural language processing (NLP), and audio processing.
- **Pattern Recognition:** Most models require machine learning engineer intervention, but deep learning models can detect all kinds of patterns automatically.

In this tutorial, we are going to dive into the world of deep learning and discover all the key concepts required for you to start a career in artificial intelligence (AI). If you're looking to

learn with some practical exercises, check out our course, [An Introduction to Deep Learning in Python](#).

Core Concepts of Deep Learning

Before diving into the intricacies of deep learning algorithms and their applications, it's essential to understand the foundational concepts that make this technology so revolutionary. This section will introduce you to the building blocks of deep learning: neural networks, deep neural networks, and activation functions.

Neural networks

At the heart of deep learning are neural networks, which are computational models inspired by the human brain. These networks consist of interconnected nodes, or "neurons," that work together to process information and make decisions. Just like our brain has different regions for different tasks, a neural network has layers designated for specific functions.

We have a full guide, [What are Neural Networks](#), which covers the essentials in more detail.

Deep neural networks

What makes a neural network "deep" is the number of layers it has between the input and output. A deep neural network has multiple layers, allowing it to learn more complex features and make more accurate predictions. The "depth" of these networks is what gives deep learning its name and its power to solve intricate problems.

Our [introduction to deep neural networks](#) tutorial covers the significance of DNNs in deep learning and artificial intelligence.

Activation functions

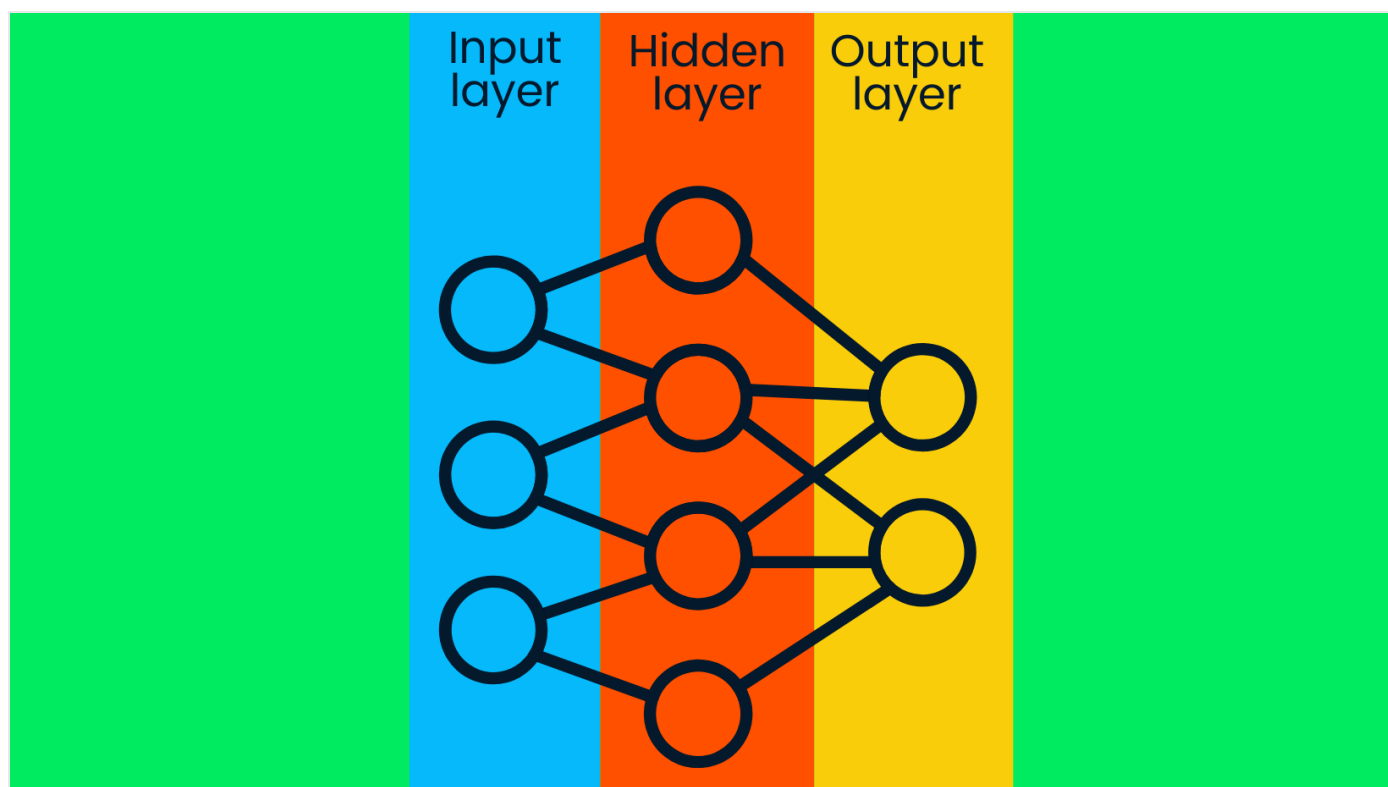
In a neural network, activation functions are like the decision-makers. They determine what information should be passed along to the next layer. These functions add a level of complexity, enabling the network to learn from the data and make nuanced decisions.

How Deep Learning Works

Deep learning uses feature extraction to recognize similar features of the same label and then uses [decision boundaries](#) to determine which features accurately represent each label. In the cats and dogs classification, the deep learning models will extract information such as the eyes, face, and body shape of animals and divide them into two classes.

The deep learning model consists of deep neural networks. The simple neural network consists of an input layer, a hidden layer, and an output layer. Deep learning models consist

of multiple hidden layers, with additional layers that the model's accuracy has improved.

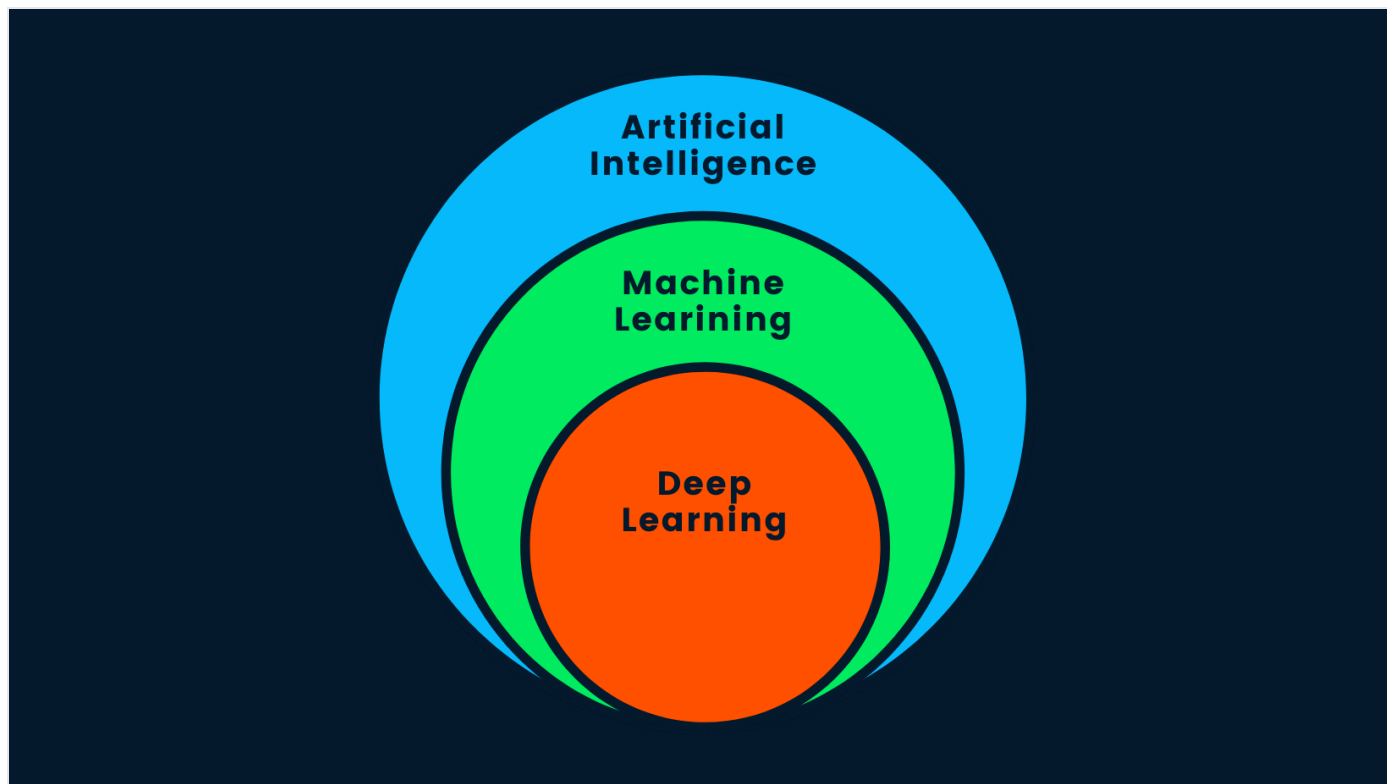


Simple Neural Network

The input layers contain raw data, and they transfer the data to hidden layers' nodes. The hidden layers' nodes classify the data points based on the broader target information, and with every subsequent layer, the scope of the target value narrows down to produce accurate assumptions. The output layer uses hidden layer information to select the most probable label. In our case, accurately predicting a dog's image rather than a cat's.

Artificial Intelligence vs. Deep Learning

Let's answer one of the most frequently asked questions on the internet: "Is deep learning artificial intelligence?". The short answer is yes. Deep learning is a subset of machine learning, and machine learning is a subset of AI.



AI vs. ML vs. DL

Artificial intelligence is the **concept that intelligent machines can be built to mimic human behavior** or surpass human intelligence. AI uses machine learning and deep learning methods to complete human tasks. In short, AI is deep learning as it is the most advanced algorithm capable of making intelligent decisions.

What is Deep Learning Used For?

Recently, the world of technology has seen a surge in artificial intelligence applications, and they all are powered by deep learning models. The applications range from recommending movies on Netflix to Amazon warehouse management systems.

In this section, we are going to learn about some of the most famous applications built using deep learning. This will help you realize the full potential of deep neural networks.

Computer Vision

Computer vision (CV) is used in self-driving cars to detect objects and avoid collisions. It is also used for face recognition, pose estimation, image classification, and anomaly detection.



Face Recognition

Automatic Speech Recognition

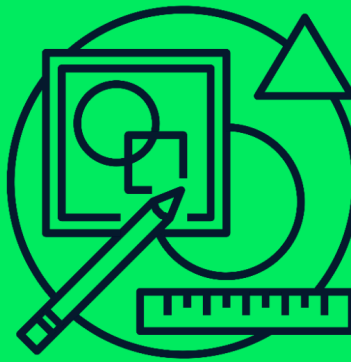
Automatic speech recognition (ASR) is used by billions of people worldwide. It is in our phones and is commonly activated by saying "Hey, Google" or "Hi, Siri." Such audio applications are also used for text-to-speech, audio classification, and voice activity detection.



Speech Pattern Recognition

Generative AI

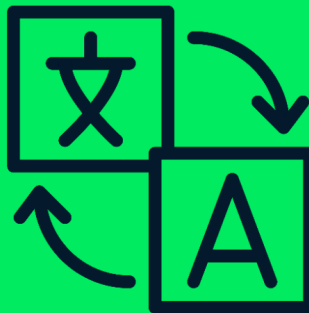
Generative AI has seen a surge in demand as **CryptoPunk NFT just sold for \$1 million**. CryptoPunk is a generative art collection that was created using deep learning models. The introduction of the **GPT-4** model by **OpenAI** has revolutionized the text generation domain with its powerful ChatGPT tool; now, you can teach models to write an entire novel or even write code for your **data science projects**.



Generative Art

Translation

Deep learning translation is not limited to language translation, as we are now able to translate photos to text by using [OCR](#), or translate text to images by using [NVIDIA GauGAN2](#).



Language Translation

Time Series Forecast

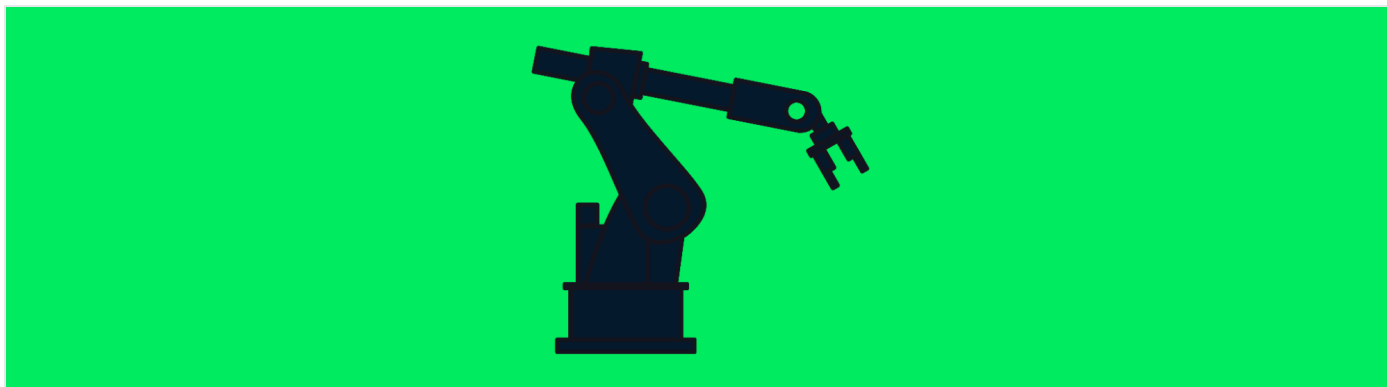
[Time series forecasting](#) is used for predicting market crashes, stock prices, and changes in the weather. The financial sector survives on speculation and future projections. Deep learning and time series models are better than humans in detecting patterns and so are pivotal tools in this and similar industries.



Time Series Forecast

Automation

Deep learning is used for automating tasks, for example, training robots for warehouse management. The most popular application is playing video games and getting better at solving puzzles. Recently, [OpenAI's Dota AI beat pro team OG](#), which shocked the world as people were not expecting all five bots to outsmart the world champions.



Robotic Arm Powered by Reinforcement Learning

Customer Feedback

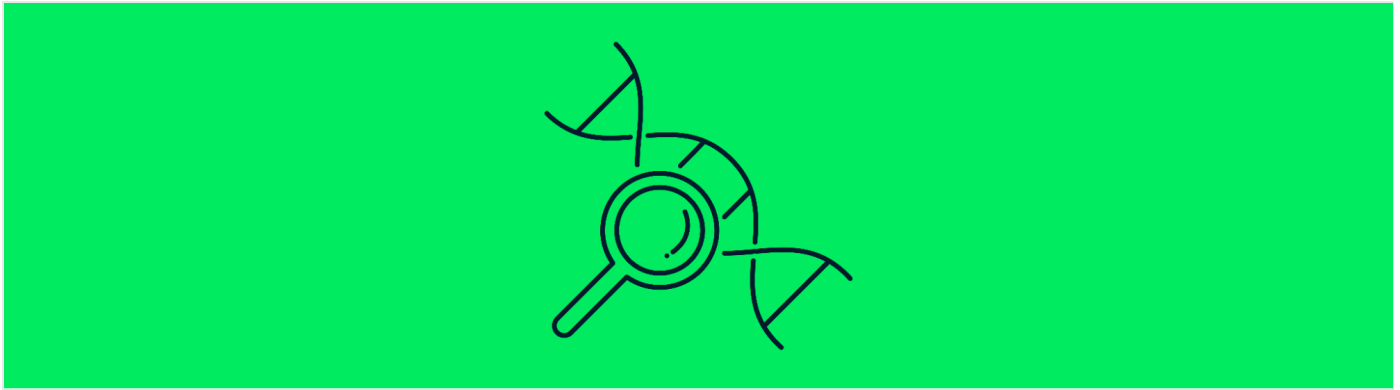
Deep learning is used for [handling customers' feedback and complaints](#). It is used in every chatbot application to provide seamless customer services.



Customer Feedback

Biomedical

This field has benefited the most with the introduction of deep learning. **DL is used in biomedicine** to detect cancer, build stable medicine, for anomaly detection in chest X-rays, and to assist medical equipment.



Analyzing DNA Sequences

Deep Learning Models

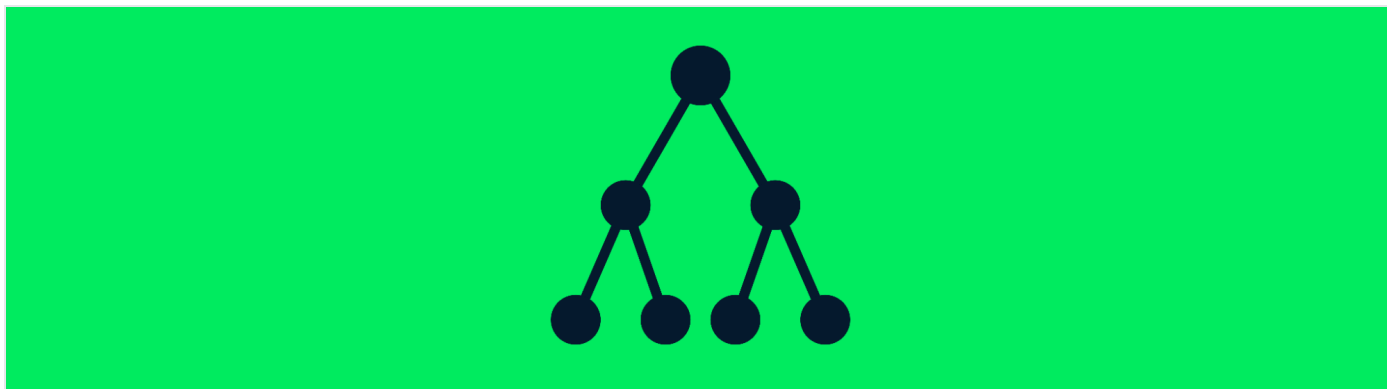
Let's learn about different types of deep learning models and how they work.

Supervised Learning

Supervised learning uses a labeled dataset to train models to either classify data or predict values. The dataset contains features and target labels, which allow the algorithm to learn over time by minimizing the loss between predicted and actual labels. Supervised learning can be divided into classification and regression problems.

Classification

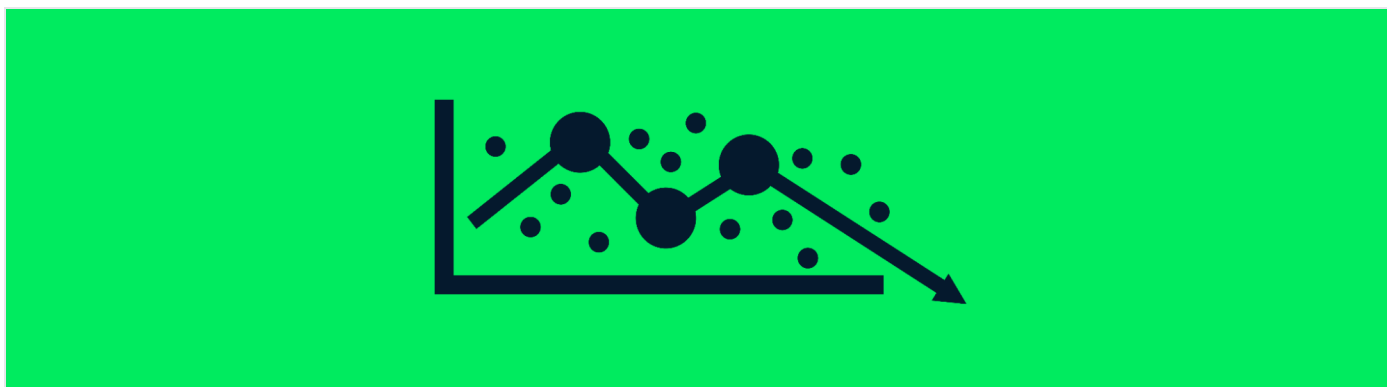
The classification algorithm divides the dataset into various categories based on feature extractions. The popular deep learning models are **ResNet50** for image classification and **BERT (language model)** for text classification.



Classification

Regression

Instead of dividing the dataset into categories, the regression model learns the relationship between input and output variables to predict the outcome. Regression models are commonly used for predictive analysis, weather forecasting, and predicting stock market performance. **LSTM** and **RNN** are popular deep learning regression models.

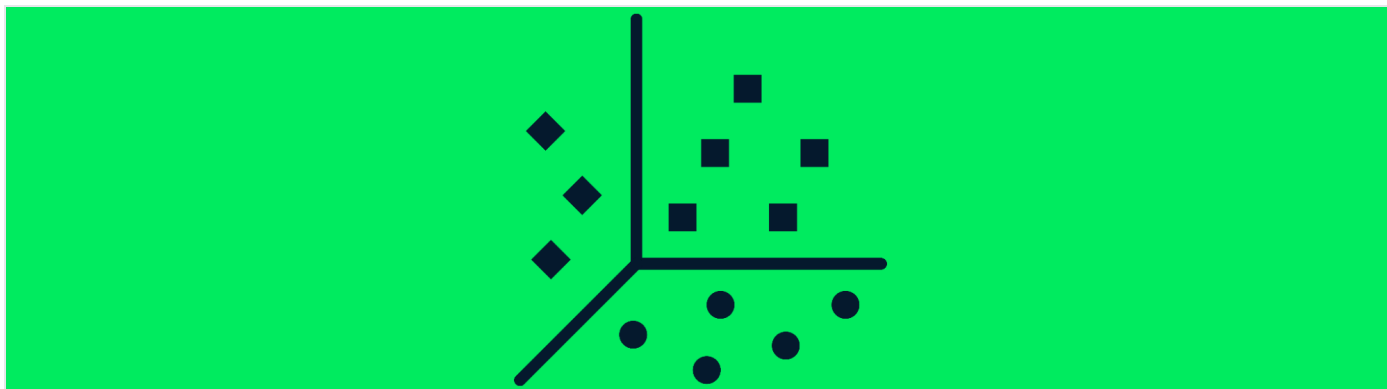


Linear Regression

Unsupervised Learning

Unsupervised learning algorithms learn the pattern within an unlabeled dataset and create clusters. Deep learning models can learn hidden patterns without human intervention and these models are often used in recommendation engines.

Unsupervised learning is used for grouping various species, medical imaging, and market research. The most common deep learning model for clustering is the **deep embedded clustering algorithm**.

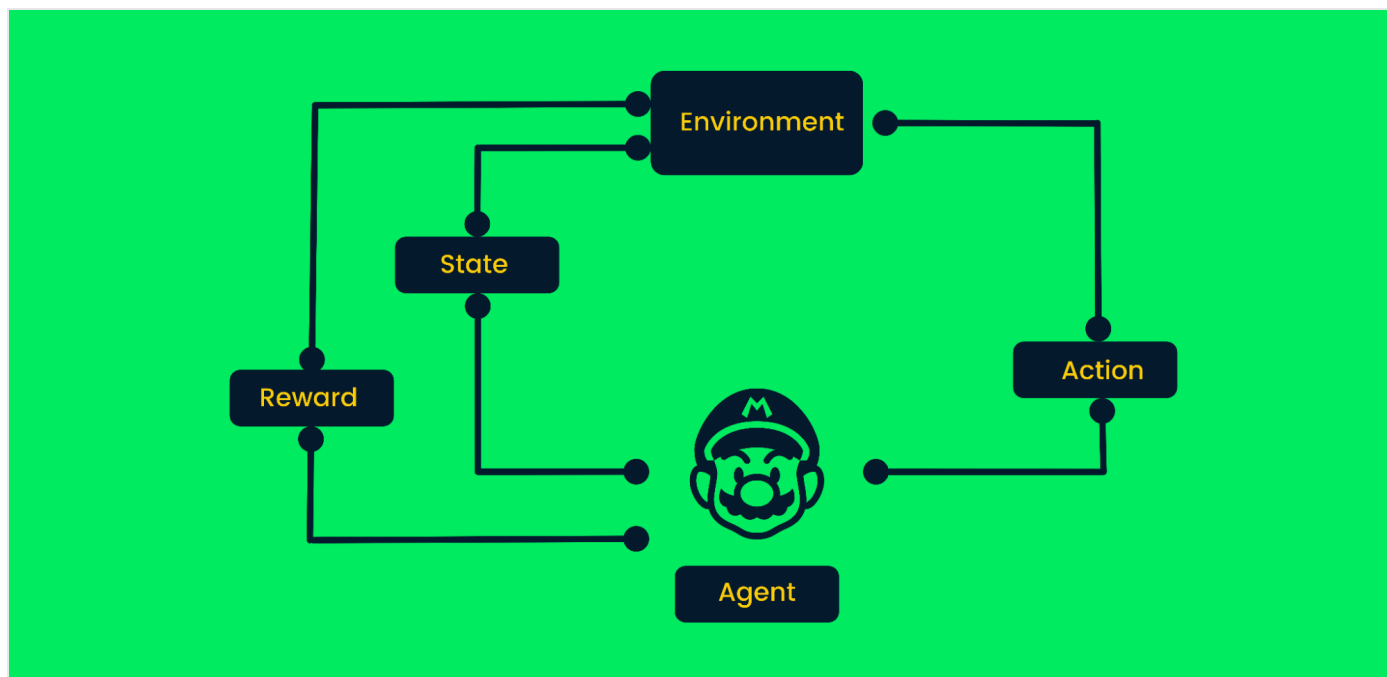


Clustering of Data

Reinforcement Learning

Reinforcement learning (RL) is a machine learning method where agents learn various behaviors from the environment. This agent takes random actions and gets rewards. The agent learns to achieve goals by trial and error in a complex environment without human intervention.

Just like a baby with encouragement from its parents learns to walk, the AI learns to perform certain tasks by maximizing rewards, and the designer sets the rewards policy. Recently, RL has seen high demands in automation due to advancements in robotics, self-driving cars, defeating pro players in games, and landing rockets back to earth.



Reinforcement Learning Framework

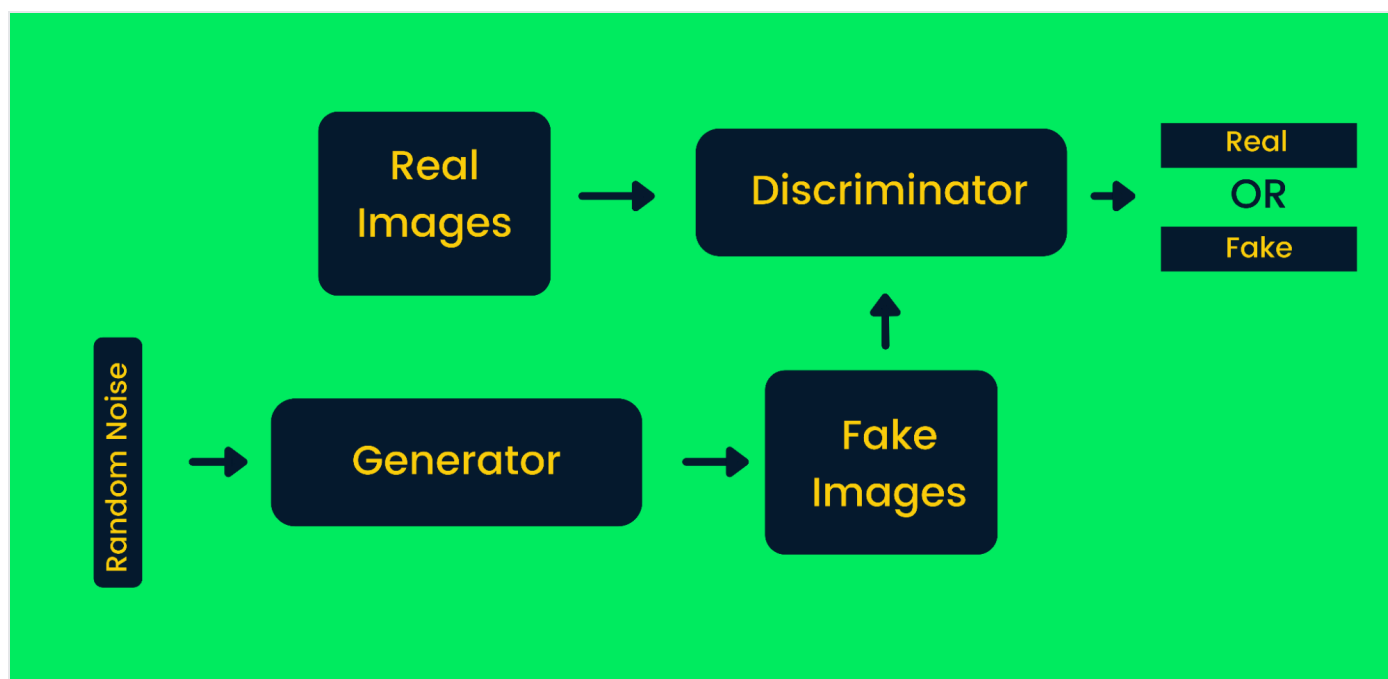
Let's take the Mario video game as an example:

- At the start, the **agent** (Mario character) receives **state** zero from the **environment**.
- Based on the state, an agent will take an **action**, in our case, Mario moved right.
- Now the state has changed, and the character is in a new frame.
- The agent receives a **reward**, as by moving right the character is not dead. Our main goal is to maximize rewards.

The agent will continue the loop of taking action and maximizing the rewards until it reaches the end of the stage or dies. Learn more at [An Introduction to Reinforcement Learning](#).

Generative Adversarial Networks

Generative adversarial networks (GANs) use two neural networks, and together, they produce synthetic instances of original data. GANs have gained a lot of popularity in recent years as they are able to mimic some of the great artists to produce masterpieces. They are widely used for generating synthetic art, video, music, and texts. Learn more about real work applications at [Generative Adversarial Networks Tutorial](#).



Generative Adversarial Network Framework

How GANs work in synthetic image generation:

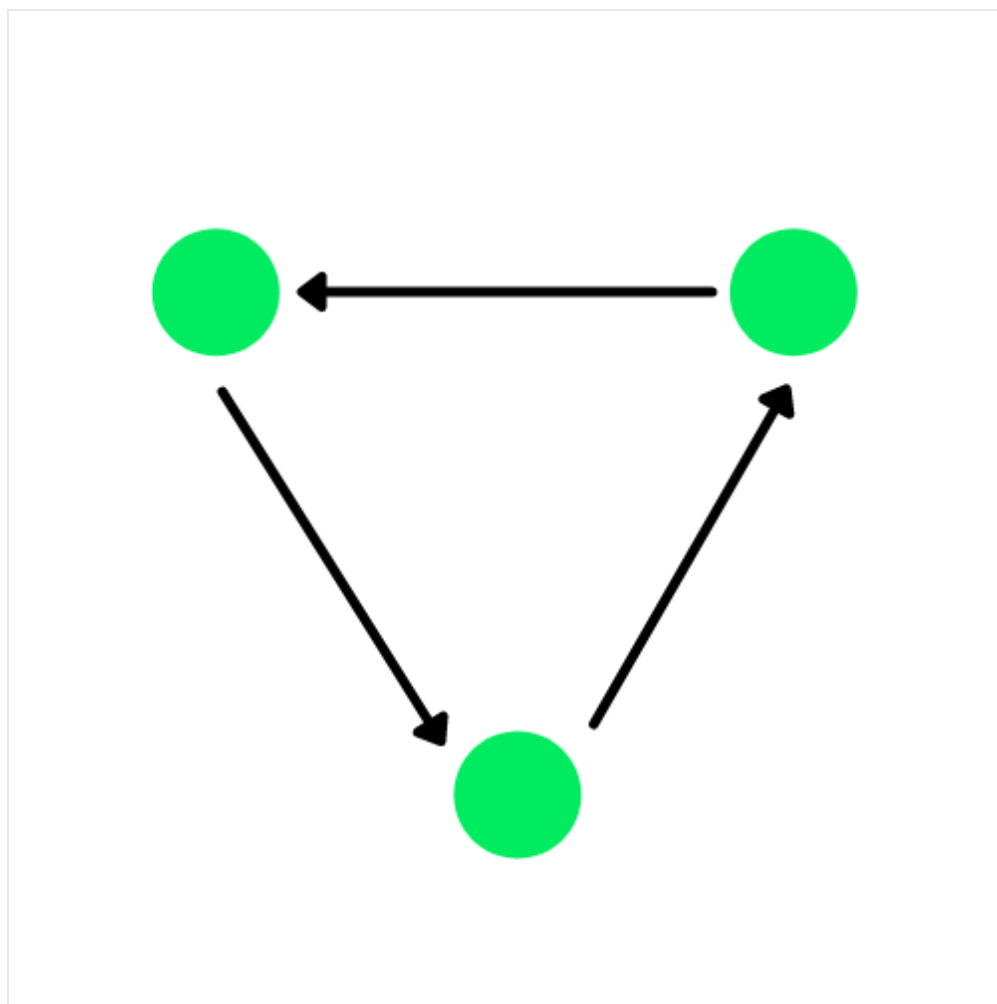
- First, the **generator** networks take input of random noise and generate fake images.
- The generated and the real images are fed into the discriminator.

- The **discriminator** decides whether the generated image is real or not. It returns probabilities from zero to one, where zero represents a fake image, and one represents an authentic image.

The GANs' architecture contains two feedback loops. The discriminator is in a feedback loop with real images, whereas the generator is in a feedback loop with a discriminator. They work in sync to produce more authentic images.

Graph Neural Network

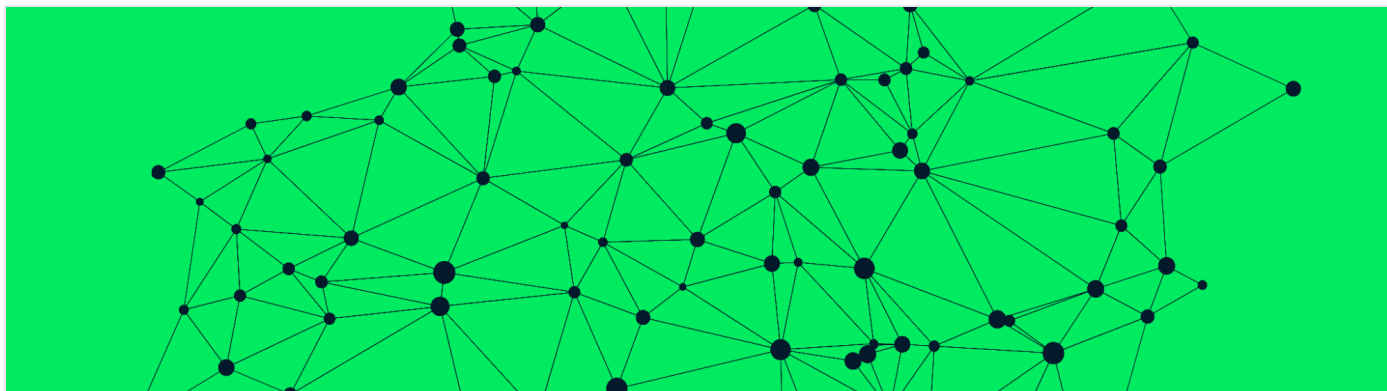
A graph is a data structure that consists of **edges** and **vertices**. The edges can be directed if there are directional dependencies between vertices (**nodes**), also known as **directed graphs**. The green circles in the diagram below are nodes, and the arrows represent the edges.



A Directed Graph

A **graph neural network** (GNN) is a type of deep learning architecture that directly operates on graph structures. GNNs are applied in large dataset analysis, recommendation systems,

and computer visions.



A Graph Network

They are also used for node classification, link prediction, and clustering. In some cases, **graph neural networks** have performed better than convolution neural networks, for example when recognizing objects and predicting semantic relations.

Natural Language Processing

Natural language processing (NLP) uses deep learning technology to **aid computers to learn a natural human language**. NLP uses deep learning to read, decipher, and understand human language. It is widely used for processing speech, text, and images. The introduction of **transfer learning** has taken NLP to the next level as we are able to fine-tune the model with a few samples and achieve state-of-the-art performance.



Subcategories of NLP

NLP can be divided into multiple fields:

- **Translation:** translating languages, molecular structure, and mathematical equations
- **Summarization:** summarizing large chunks of text into a few lines while maintaining the key information.

- **Classification:** dividing the text into various categories.
- **Generation:** text to text generation; it can be used to generate entire essays with a single line of text.
- **Conversational:** Virtual assistant, retaining past knowledge of conversation and mimicking human conversations.
- **Answering questions:** AI answers questions by using Q&A data.
- **Feature extraction:** to detect patterns in text or extract information such as "name entity recognition" and "part of speech".
- **Sentence similarities:** evaluating similarities between various texts.
- **Text to speech:** converting text to audible speech.
- **Automatic speech recognition:** understanding various sounds and converting them into text.
- **Optical character recognition:** extraction of text data from images.

If you want to test all the various applications of NLP, try [Hugging Face Spaces](#). The Spaces hosts all types of web applications that you can play around with to get inspiration for your NLP project.

A Deeper Look at Deep Learning Concepts

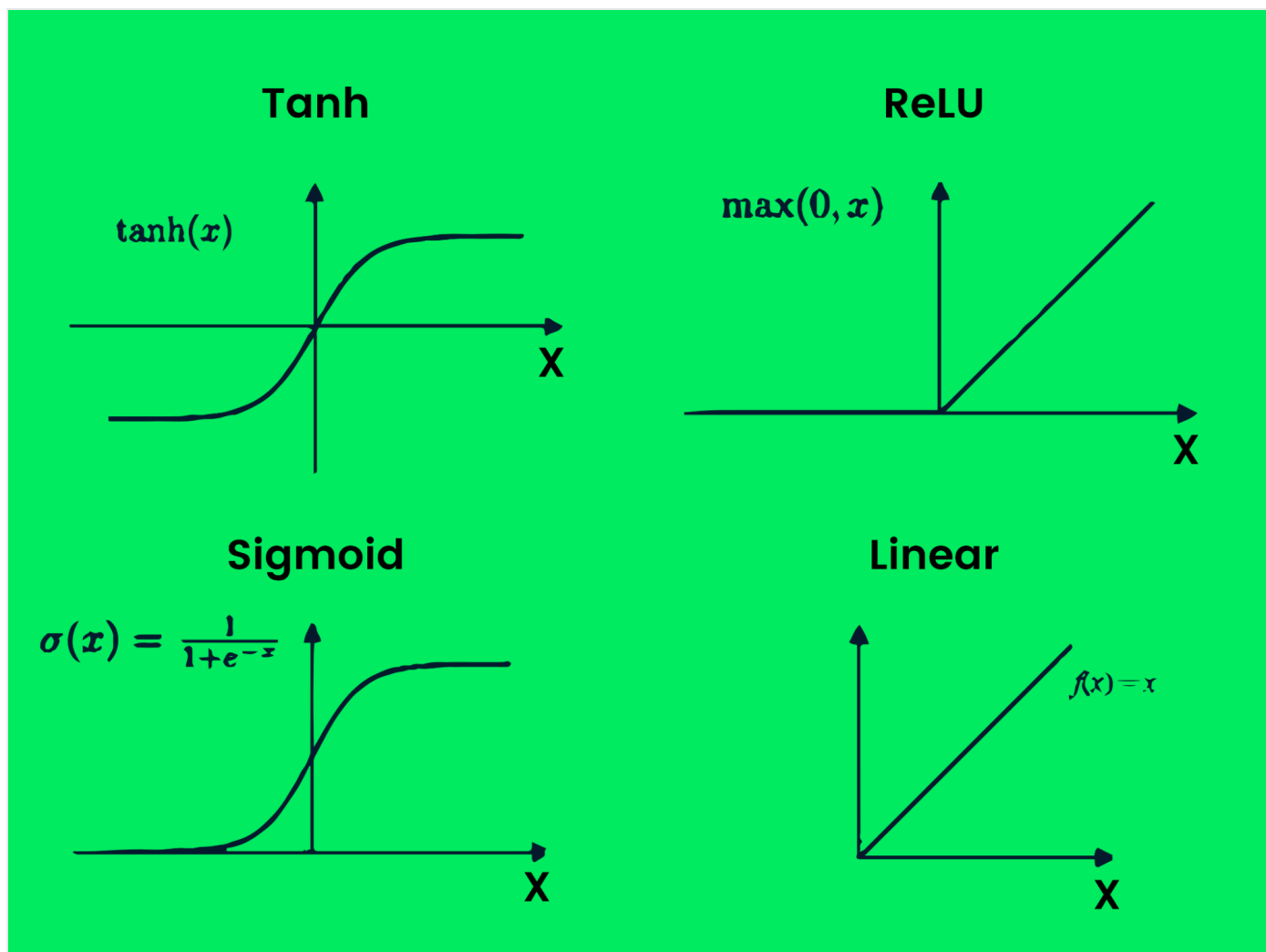
Activation Functions

In neural networks, the activation function produces output decision boundaries and is used to improve performance of the model. The activation function is a mathematical expression that decides whether the input should pass through a neuron or not based on its significance. It also provides non-linearity to networks. Without activation function, the neural network becomes a simple linear regression model.

There are several types of activation functions:

- Tanh
- ReLU
- Sigmoid
- Linear
- Softmax

- Swish



Activation Function

These functions produce various output boundaries, as shown in the image above. With multiple layers and activation functions, you can solve any complex problem. Learn more about [what are activation functions in deep learning?](#)

Loss Function

The loss function is the difference between actual and predicted values. It allows neural networks to track the model's overall performance. Depending on specific problems, we chose a certain type of function, for example, **mean squared error**.

$$\text{Loss} = \text{Sum} (\text{Predicted} - \text{Actual})^2$$

The most used loss functions in deep learning are:

- Binary cross-entropy
- Categorical hinge

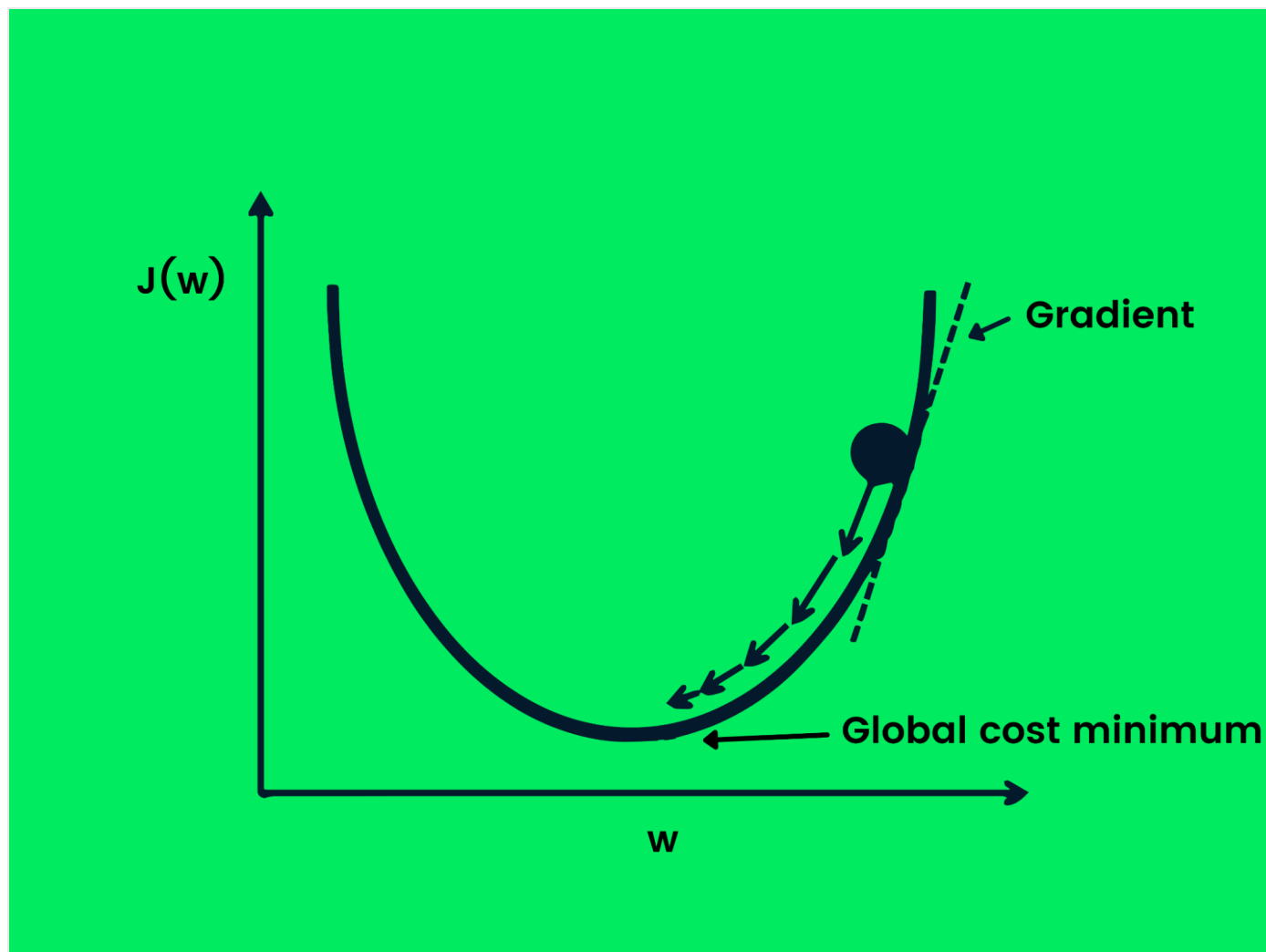
- Mean squared error
- Huber
- Sparse categorical cross-entropy

Backpropagation

In **forwarding propagation**, we initialize our neural network with random inputs to produce an output that is random too. To make our model perform better, we adjust weights randomly using **backpropagation**. To track the model's performance, we need a loss function that will find **global minima** to maximize the model's accuracy.

Stochastic Gradient Descent

Gradient descent is used to optimize loss function by changing weights in a controlled way to achieve minimum loss. Now we have an objective, but we need direction on whether to increase or decrease the **weights** to achieve better performance. The derivative of the loss function will give us direction and we can use it to update the weights of the network.



Gradient Descent

The equation below shows how weights are updated using gradient descent.

$$\mathbf{w} = \mathbf{w} - \mathbf{J}\mathbf{w}$$

In **stochastic gradient descent**, samples are divided into batches instead of using the entire dataset to optimize gradient descent. This is useful if you want to achieve minimum loss faster and optimize computational power.

Hyperparameter

Hyperparameters are the tunable parameters adjusted before running the training process. These parameters directly affect model performance and help you achieve faster global minima.

List of most used hyperparameters:

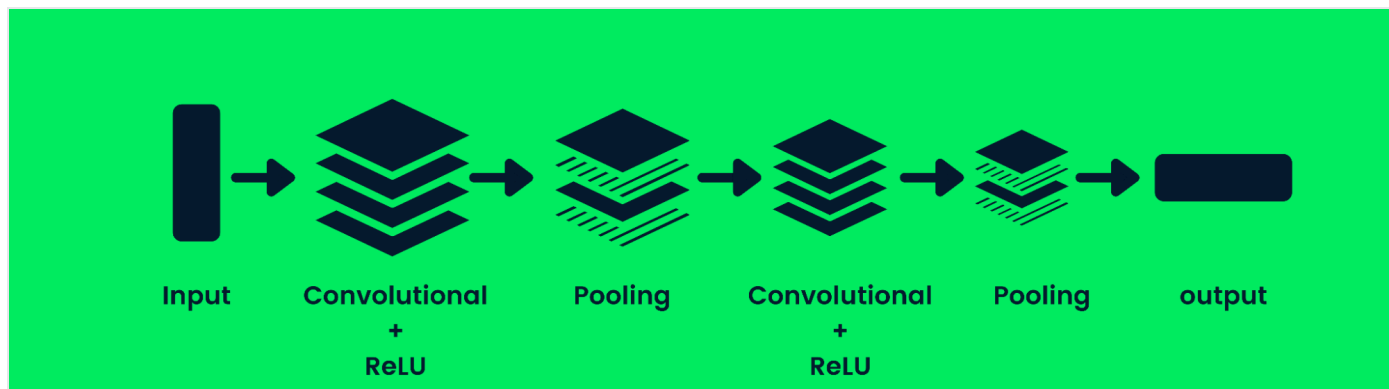
- **Learning rate:** step size of each iteration and can be set from 0.1 to 0.0001. In short, it determines the speed at which the model learns.
- **Batch size:** number of samples passed through a neural network at a time.
- **Number of epochs:** an iteration of how many times the model changes weights. Too many epochs can cause models to overfit and too few can cause models to underfit, so we have to choose a medium number.

Learn more about how these components work together by following [Keras Tutorial: Deep Learning in Python](#).

Popular Algorithms

Convolutional Neural Networks

The convolutional neural network (CNN) is a **feed-forward neural network** capable of processing a structured array of data. It is widely used for computer vision applications such as image classification.

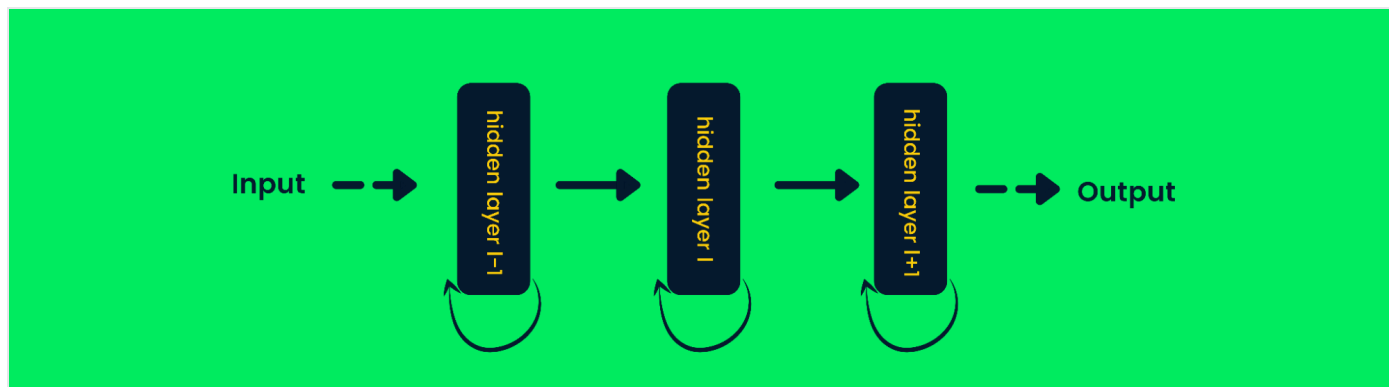


Convolution Neural Network Architecture

CNNs are good at recognizing patterns, lines, and shapes. The CNN consists of a **convolutional** layer, **pooling** layer, and output layer (fully connected layers). The image classification models usually contain multiple convolution layers, followed by pooling layers, as additional layers increase the accuracy of the model. Learn more about convolutional layers here: [Convolutional Neural Networks in Python](#).

Recurrent Neural Networks

Recurrent neural networks (RNN) are different from feed-forward networks as the output of the layer is fed back into the input to predict the output of the layer. This helps it perform better with sequential data as it can store the information of previous samples to predict future samples. Learn more at [Recurrent Neural Network \(RNN\) Tutorial: Types & Examples](#).

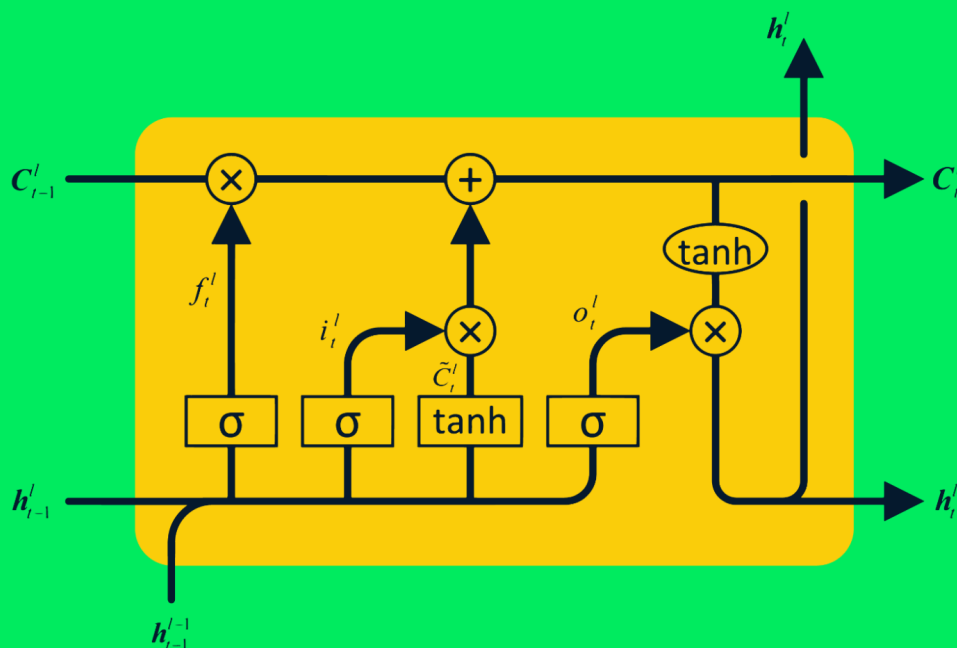


Recurrent Neural Network Architecture

In traditional neural networks, the output of the layers is calculated based on the current input values, but in RNN the output is calculated based on previous inputs too. This makes it quite good at predicting the next word, forecasting stock prices, in AI chatbots, and anomaly detection.

Long Short-term Memory Networks

Long short-term memory networks (LSTM) are advanced types of recurrent neural networks that can retain greater information on past values. It solves vanishing gradient problems that exist in simple RNN.



LSTM Architecture

The typical RNN consists of repeating neural networks with a single tanh layer, whereas LSTM consists of four interactive layers that communicate to process large sequences of data.

You can get hands-on experience with the following [Tutorial: LSTM for stock predictions](#) or the [advanced deep learning with Keras](#) course if you want to learn more about deep learning models.

Deep Learning Frameworks

There are multiple deep learning frameworks, such as MxNet, CNTK, and Caffe2, but we will be learning about the most popular frameworks.

Tensorflow

Tensorflow (TF) is an open-source library used for creating deep learning applications. It includes all the necessary tools for you to experiment and develop commercial AI products. It supports both CPU, GPU, and TPU for training complex models. TF was originally developed by the Google AI team for internal use and is now available to the public.

The Tensorflow API is available for browser-based applications, mobile devices, and TensorFlow Extended is ideal for production. TF has now become the industry standard, and it is used for both academic research and deploying deep learning models in production.

TF also comes with **Tensorboard**, which is a dashboard capable of analyzing your machine learning experiments. Recently, Tensorflow developers have integrated Keras into its framework, which is popular for developing deep neural networks. Learn more at the [Introduction to TensorFlow in Python Course](#).

Keras

Keras is a neural network framework written in Python and capable of running on multiple frameworks such as Tensorflow and Theano. Keras is an open-source library developed to enable fast experimentation in deep learning so that you can easily convert your concepts into working AI applications.

The documentation is quite easy to understand, and the API is similar to **Numpy**, which allows you to easily integrate it with any data science project. Just like TF, Keras can also run on CPU, GPU, and TPU, based on available hardware. Learn more at [Introduction to Deep Learning with Keras](#).

PyTorch

PyTorch is the most popular and easiest deep learning framework. It uses tensor instead of Numpy array to perform fast numerical computation powered by GPU. PyTorch is mainly used for deep learning and developing complex machine learning models.

Academic researchers prefer using PyTorch because of its flexibility and ease of use. It is written in C++ and Python, and it also comes with GPUs and TPUs acceleration. It has become a one-stop solution for all deep learning problems. If you'd like to learn more about PyTorch, try taking the [Introduction to Deep Learning with PyTorch](#) course.

Conclusion

In this tutorial, we have covered all the what deep learning is, some of the basics of deep learning, how it works, and its applications. We have also learned how deep neural networks