**MARKET BASKET INSIGHTS**

# Introduction to YOLO Algorithm for Object Detection

YOLO is an algorithm that uses neural networks to provide real-time object detection. This algorithm is popular because of its speed and accuracy. It has been used in various applications to detect traffic signals, people, parking meters, and animals.

This article introduces readers to the YOLO algorithm for object detection and explains how it works. It also highlights some of its real-life applications.

### **Introduction to object detection**

Object detection is a phenomenon in [computer vision](https://www.section.io/engineering-education/computer-vision-straight-lines/) that involves the detection of various objects in digital images or videos. Some of the objects detected include people, cars, chairs, stones, buildings, and animals.

This phenomenon seeks to answer two basic questions:

1. What is the object? This question seeks to identify the object in a specific image.
2. Where is it? This question seeks to establish the exact location of the object within the image.

Object detection consists of various approaches such as [fast R-CNN](https://towardsdatascience.com/understanding-fast-r-cnn-and-faster-r-cnn-for-object-detection-adbb55653d97?gi=fea1a85170b6), [Retina-Net](https://developers.arcgis.com/python/guide/how-retinanet-works/), and [Single-Shot MultiBox Detector (SSD)](https://iq.opengenus.org/single-shot-detection-ssd-algorithm/). Although these approaches have solved the challenges of data limitation and modeling in object detection, they are not able to detect objects in a single algorithm run. **YOLO algorithm** has gained popularity because of its superior performance over the aforementioned object detection techniques.

### **What is YOLO?**

YOLO is an abbreviation for the term ‘You Only Look Once’. This is an algorithm that detects and recognizes various objects in a picture (in real-time). Object detection in YOLO is done as a regression problem and provides the class probabilities of the detected images.

YOLO algorithm employs convolutional neural networks (CNN) to detect objects in real-time. As the name suggests, the algorithm requires only a single forward propagation through a neural network to detect objects.

This means that prediction in the entire image is done in a single algorithm run. The CNN is used to predict various class probabilities and bounding boxes simultaneously.

The YOLO algorithm consists of various variants. Some of the common ones include tiny YOLO and YOLOv3.

### **Why the YOLO algorithm is important**

YOLO algorithm is important because of the following reasons:

* **Speed:** This algorithm improves the speed of detection because it can predict objects in real-time.
* **High accuracy:** YOLO is a predictive technique that provides accurate results with minimal background errors.
* **Learning capabilities:** The algorithm has excellent learning capabilities that enable it to learn the representations of objects and apply them in object detection.

### **Applications of YOLO**

YOLO algorithm can be applied in the following fields:

* **Autonomous driving:** YOLO algorithm can be used in autonomous cars to detect objects around cars such as vehicles, people, and parking signals. Object detection in autonomous cars is done to avoid collision since no human driver is controlling the car.
* **Wildlife:** This algorithm is used to detect various types of animals in forests. This type of detection is used by wildlife rangers and journalists to identify animals in videos (both recorded and real-time) and images. Some of the animals that can be detected include giraffes, elephants, and bears.
* **Security:** YOLO can also be used in security systems to enforce security in an area. Let’s assume that people have been restricted from passing through a certain area for security reasons. If someone passes through the restricted area, the YOLO algorithm will detect him/her, which will require the security personnel to take further action.

# Introduction to YOLO Algorithm for Object Detection

##### **April 15, 2021**

YOLO is an algorithm that uses neural networks to provide real-time object detection. This algorithm is popular because of its speed and accuracy. It has been used in various applications to detect traffic signals, people, parking meters, and animals.

This article introduces readers to the YOLO algorithm for object detection and explains how it works. It also highlights some of its real-life applications.

### **Introduction to object detection**

Object detection is a phenomenon in [computer vision](https://www.section.io/engineering-education/computer-vision-straight-lines/) that involves the detection of various objects in digital images or videos. Some of the objects detected include people, cars, chairs, stones, buildings, and animals.

This phenomenon seeks to answer two basic questions:

1. What is the object? This question seeks to identify the object in a specific image.
2. Where is it? This question seeks to establish the exact location of the object within the image.

Object detection consists of various approaches such as [fast R-CNN](https://towardsdatascience.com/understanding-fast-r-cnn-and-faster-r-cnn-for-object-detection-adbb55653d97?gi=fea1a85170b6), [Retina-Net](https://developers.arcgis.com/python/guide/how-retinanet-works/), and [Single-Shot MultiBox Detector (SSD)](https://iq.opengenus.org/single-shot-detection-ssd-algorithm/). Although these approaches have solved the challenges of data limitation and modeling in object detection, they are not able to detect objects in a single algorithm run. **YOLO algorithm** has gained popularity because of its superior performance over the aforementioned object detection techniques.

### **What is YOLO?**

YOLO is an abbreviation for the term ‘You Only Look Once’. This is an algorithm that detects and recognizes various objects in a picture (in real-time). Object detection in YOLO is done as a regression problem and provides the class probabilities of the detected images.

YOLO algorithm employs convolutional neural networks (CNN) to detect objects in real-time. As the name suggests, the algorithm requires only a single forward propagation through a neural network to detect objects.

This means that prediction in the entire image is done in a single algorithm run. The CNN is used to predict various class probabilities and bounding boxes simultaneously.

The YOLO algorithm consists of various variants. Some of the common ones include tiny YOLO and YOLOv3.

### **Why the YOLO algorithm is important**

YOLO algorithm is important because of the following reasons:

* **Speed:** This algorithm improves the speed of detection because it can predict objects in real-time.
* **High accuracy:** YOLO is a predictive technique that provides accurate results with minimal background errors.
* **Learning capabilities:** The algorithm has excellent learning capabilities that enable it to learn the representations of objects and apply them in object detection.

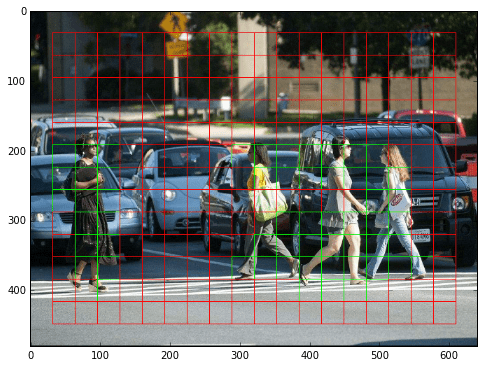
### **How the YOLO algorithm works**

YOLO algorithm works using the following three techniques:

* Residual blocks
* Bounding box regression
* Intersection Over Union (IOU)

#### **Residual blocks**

First, the image is divided into various grids. Each grid has a dimension of S x S. The following image shows how an input image is divided into grids.



[Image Source](https://www.guidetomlandai.com/assets/img/computer_vision/grid.png)

In the image above, there are many grid cells of equal dimension. Every grid cell will detect objects that appear within them. For example, if an object center appears within a certain grid cell, then this cell will be responsible for detecting it.

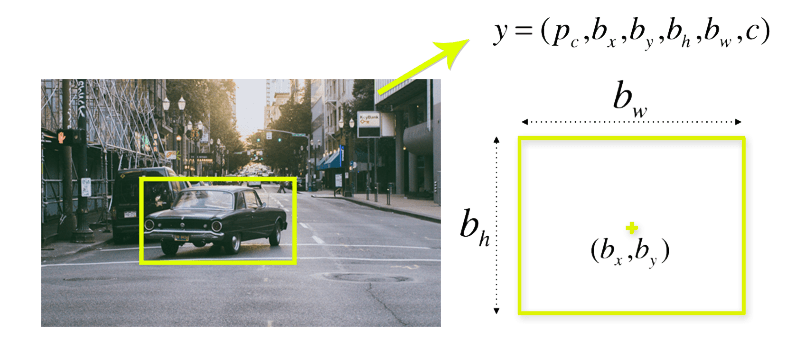
#### **Bounding box regression**

A bounding box is an outline that highlights an object in an image.

Every bounding box in the image consists of the following attributes:

* Width (bw)
* Height (bh)
* Class (for example, person, car, traffic light, etc.)- This is represented by the letter c.
* Bounding box center (bx,by)

The following image shows an example of a bounding box. The bounding box has been represented by a yellow outline.



[Image Source](https://appsilondatascience.com/assets/uploads/2018/08/bbox-1.png)

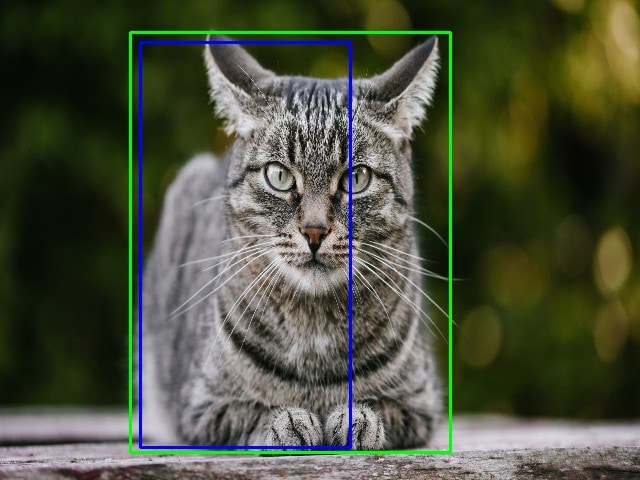
YOLO uses a single bounding box regression to predict the height, width, center, and class of objects. In the image above, represents the probability of an object appearing in the bounding box.

#### **Intersection over union (IOU)**

Intersection over union (IOU) is a phenomenon in object detection that describes how boxes overlap. YOLO uses IOU to provide an output box that surrounds the objects perfectly.

Each grid cell is responsible for predicting the bounding boxes and their confidence scores. The IOU is equal to 1 if the predicted bounding box is the same as the real box. This mechanism eliminates bounding boxes that are not equal to the real box.

The following image provides a simple example of how IOU works.

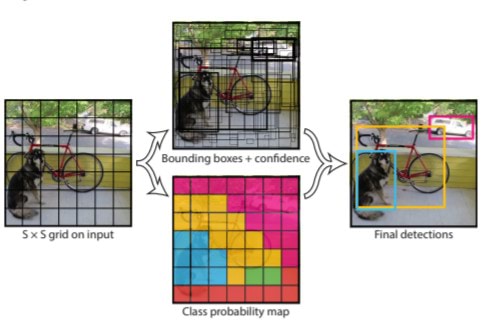


[Image Source](https://miro.medium.com/max/640/1*VuAsK1Wwa_mOxW2nK2UovQ.jpeg)

In the image above, there are two bounding boxes, one in green and the other one in blue. The blue box is the predicted box while the green box is the real box. YOLO ensures that the two bounding boxes are equal.

#### **Combination of the three techniques**

The following image shows how the three techniques are applied to produce the final detection results.



[Image Source](https://www.guidetomlandai.com/assets/img/computer_vision/YOLO.PNG)

First, the image is divided into grid cells. Each grid cell forecasts B bounding boxes and provides their confidence scores. The cells predict the class probabilities to establish the class of each object.

For example, we can notice at least three classes of objects: a car, a dog, and a bicycle. All the predictions are made simultaneously using a single convolutional neural network.

Intersection over union ensures that the predicted bounding boxes are equal to the real boxes of the objects. This phenomenon eliminates unnecessary bounding boxes that do not meet the characteristics of the objects (like height and width). The final detection will consist of unique bounding boxes that fit the objects perfectly.

For example, the car is surrounded by the pink bounding box while the bicycle is surrounded by the yellow bounding box. The dog has been highlighted using the blue bounding box.

### **Applications of YOLO**

YOLO algorithm can be applied in the following fields:

* **Autonomous driving:** YOLO algorithm can be used in autonomous cars to detect objects around cars such as vehicles, people, and parking signals. Object detection in autonomous cars is done to avoid collision since no human driver is controlling the car.
* **Wildlife:** This algorithm is used to detect various types of animals in forests. This type of detection is used by wildlife rangers and journalists to identify animals in videos (both recorded and real-time) and images. Some of the animals that can be detected include giraffes, elephants, and bears.
* **Security:** YOLO can also be used in security systems to enforce security in an area. Let’s assume that people have been restricted from passing through a certain area for security reasons. If someone passes through the restricted area, the YOLO algorithm will detect him/her, which will require the security personnel to take further action.

### **Conclusion**

This article has provided an overview of the YOLO algorithm and how it is used in object detection. This technique provides improved detection results compared to other object detection techniques such as Fast R-CNN and Retina-Net.

To summarize:

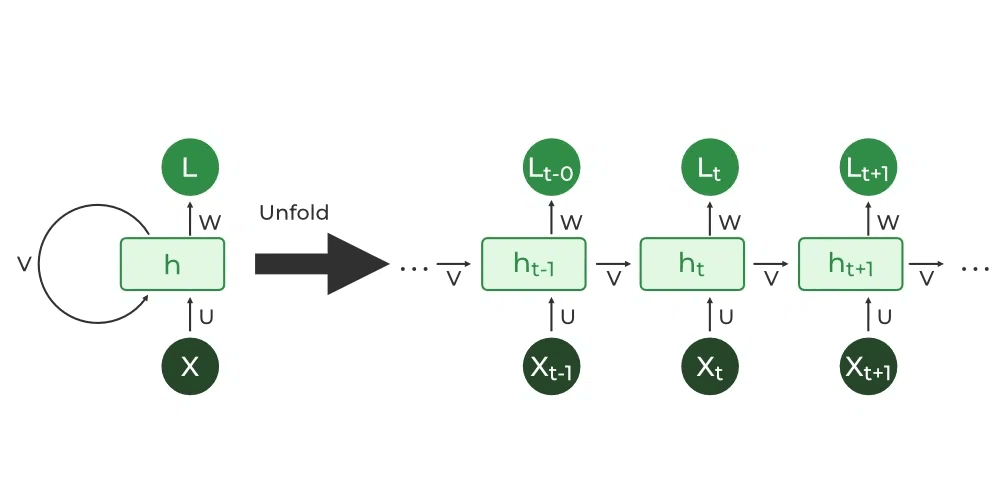
* We have gained an overview of object detection and the YOLO algorithm.
* We have gone through the main reasons why the YOLO algorithm is important.
* We have learned how the YOLO algorithm works. We have also gained an understanding of the main techniques used by YOLO to detect objects.
* We have learned the real-life applications of YOLO.

# Introduction to Recurrent Neural Network

In this article, we will introduce a new variation of neural network which is the **Recurrent Neural Network** also known as**(RNN)** that works better than a simple neural network when data is sequential like Time-Series data and text data.

## What is Recurrent Neural Network (RNN)?

Recurrent Neural Network(RNN) is a type of [Neural Network](https://www.geeksforgeeks.org/tag/neural-network/) where the output from the previous step is fed as input to the current step. In traditional neural networks, all the inputs and outputs are independent of each other, but in cases when it is required to predict the next word of a sentence, the previous words are required and hence there is a need to remember the previous words. Thus RNN came into existence, which solved this issue with the help of a Hidden Layer. The main and most important feature of RNN is its **Hidden state**, which remembers some information about a sequence. The state is also referred to as *Memory State*since it remembers the previous input to the network. It uses the same parameters for each input as it performs the same task on all the inputs or hidden layers to produce the output. This reduces the complexity of parameters, unlike other neural networks.



## Architecture Of Recurrent Neural Network

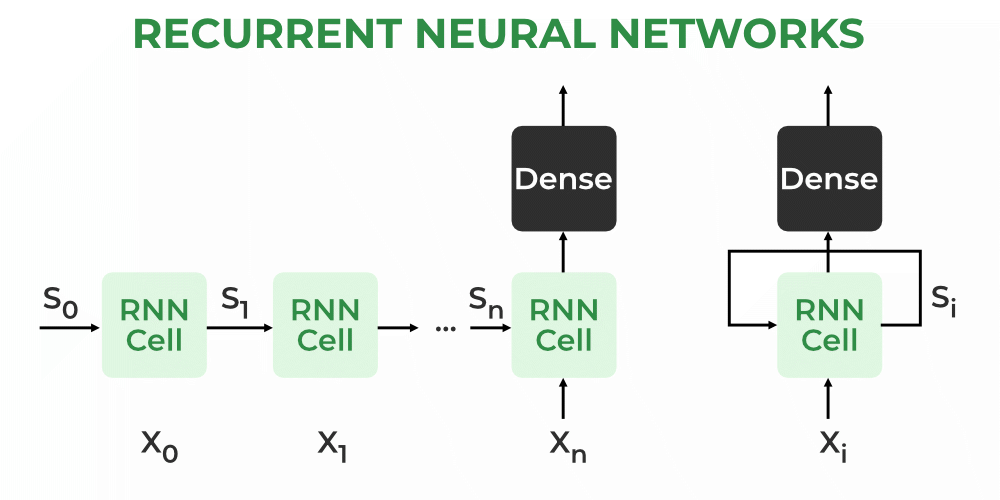
RNNs have the same input and output architecture as any other deep neural architecture. However, differences arise in the way information flows from input to output. Unlike Deep neural networks where we have different weight matrices for each Dense network in RNN, the weight across the network remains the same. It calculates state hidden state  H**i**for every input **Xi .By using the following formulas:**

**h= σ(UX + Wh-1 + B)**

**Y = O(Vh + C) Hence**

**Y = f (X, h , W, U, V, B, C)**

Here S is the State matrix which has element si as the state of the network at timestep i  
The parameters in the network are W, U, V, c, b which are shared across timestep



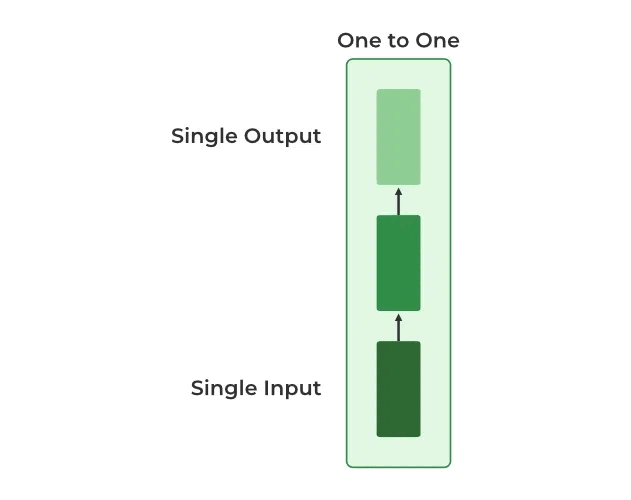
## Types Of RNN

There are four types of RNNs based on the number of inputs and outputs in the network.

1. One to One
2. One to Many
3. Many to One
4. Many to Many

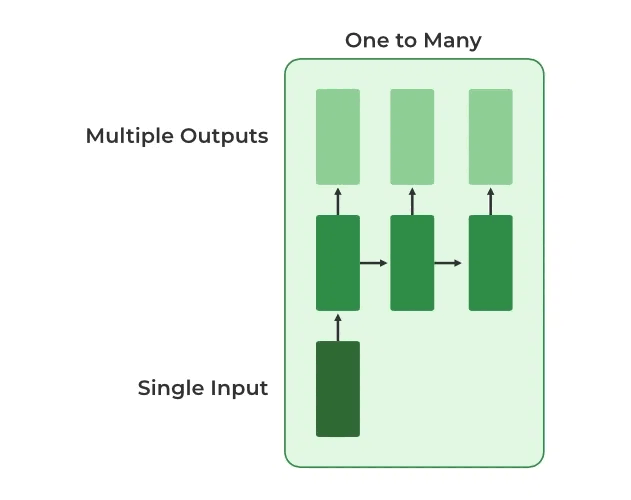
### One to One

This type of RNN behaves the same as any simple Neural network it is also known as Vanilla Neural Network. In this Neural network, there is only one input and one output.



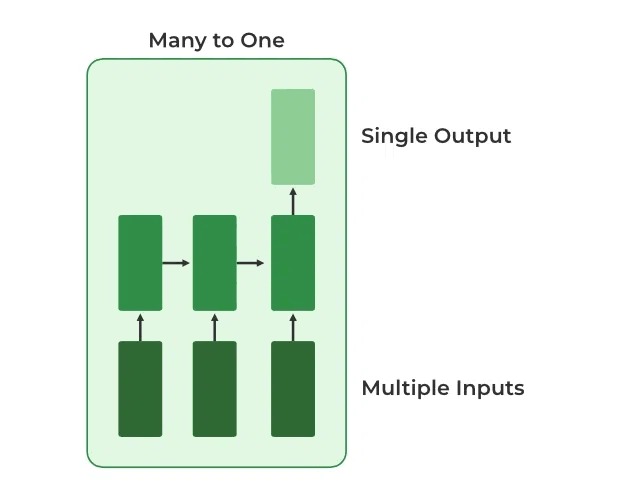
### One To Many

In this type of RNN, there is one input and many outputs associated with it. One of the most used examples of this network is Image captioning where given an image we predict a sentence having Multiple words.



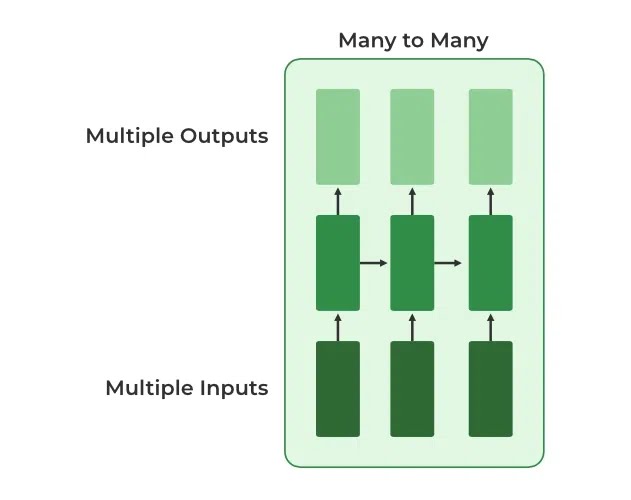
### Many to One

In this type of network, Many inputs are fed to the network at several states of the network generating only one output. This type of network is used in the problems like sentimental analysis. Where we give multiple words as input and predict only the sentiment of the sentence as output.



### Many to Many

In this type of neural network, there are multiple inputs and multiple outputs corresponding to a problem. One Example of this Problem will be language translation. In language translation, we provide multiple words from one language as input and predict multiple words from the second language as output.



## Variation Of Recurrent Neural Network (RNN)

To overcome the problems like vanishing gradient and exploding gradient descent several new advanced versions of RNNs are formed some of these are as ;

1. Bidirectional Neural Network (BiNN)
2. Long Short-Term Memory (LSTM)

**Bidirectional Neural Network (BiNN)**

A BiNN is a variation of a Recurrent Neural Network in which the input information flows in both direction and then the output of both direction are combined to produce the input. BiNN is useful in situations when the context of the input is more important such as Nlp tasks and Time-series analysis problems.

**Long Short-Term Memory (LSTM)**

[Long Short-Term Memory](https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory/amp/) works on the read-write-and-forget principle where given the input information network reads and writes the most useful information from the data and it forgets about the information which is not important in predicting the output. For doing this three new gates are introduced in the RNN. In this way, only the selected information is passed through the network.

# **Natural Language Processing**

### What do you mean by Natural Language Processing?

Natural Language Processing (NLP) is a branch of artificial intelligence (AI) that concentrates on the interaction between human language and computers.

### What is the purpose of Natural Language Processing (NLP)?

NLP involves the development of algorithms, models, and techniques that enable computers to comprehend, interpret, and generate human language in a meaningful and useful manner. It encompasses a wide array of tasks and applications related to language understanding and generation.

### How does natural language processing function?

NLP utilizes diverse techniques such as statistical modeling, machine learning, deep learning, and rule-based linguistic approaches. It includes preprocessing and analyzing textual data, constructing language models, and employing algorithms to extract insights and perform language-related tasks.

### What is the objective of NLP?

The aim of NLP is to bridge the gap between human language and computers, allowing computers to effectively understand, process, and generate natural language. NLP finds applications in various domains, including customer support, content analysis, information retrieval, virtual assistants, language translation, and more.

### How is NLP utilized in social media?

Natural Language Processing (NLP) plays a crucial role in several aspects of social media. Here are some primary applications of NLP in the realm of social media:

* Sentiment analysis

NLP techniques are employed to analyze the sentiment expressed in social media posts, comments, and reviews. This helps businesses comprehend user opinions and emotions towards their products, services, or brands. Sentiment analysis enables organizations to monitor customer satisfaction, identify potential issues, and promptly respond to customer feedback.

* Text classification and topic modeling

NLP algorithms are utilized to categorize and classify social media content into different topics or themes. This enables businesses to understand the main subjects of discussion, track trends, and identify popular topics within their industry. Text classification and topic modeling assist organizations in tailoring their content strategies, targeting specific audience segments, and engaging in relevant conversations.

* Named entity recognition

NLP techniques like named entity recognition are used to identify and extract important entities such as people, organizations, locations, and products mentioned in social media posts. This aids in understanding the context, identifying influencers or brand mentions, and tracking the reach of campaigns or events.

* Language generation

NLP models, such as ChatGPT, can generate text that resembles human language. This text can be used to compose social media captions, tweets, or responses to user queries. Language generation models assist in creating engaging and creative content, automating parts of the content creation process for social media platforms.

* Social network analysis

NLP is employed to analyze the connections and interactions between users on social media platforms. By examining the content of posts, comments, and messages, as well as network structures, NLP can help identify communities, influencers, or key users within a social network. This information can be utilized for targeted marketing, influencer identification, and relationship-building strategies.

NLP techniques provide valuable insights, automation, and enhanced user experiences, enabling businesses to effectively leverage the power of social media data.

NLP tasks

Human language is filled with ambiguities that make it incredibly difficult to write software that accurately determines the intended meaning of text or voice data. Homonyms, homophones, sarcasm, idioms, metaphors, grammar and usage exceptions, variations in sentence structure—these just a few of the irregularities of human language that take humans years to learn, but that programmers must teach natural language-driven applications to recognize and understand accurately from the start, if those applications are going to be useful.

Several NLP tasks break down human text and voice data in ways that help the computer make sense of what it's ingesting. Some of these tasks include the following:

* **Speech recognition**, also called speech-to-text, is the task of reliably converting voice data into text data. Speech recognition is required for any application that follows voice commands or answers spoken questions. What makes speech recognition especially challenging is the way people talk—quickly, slurring words together, with varying emphasis and intonation, in different accents, and often using incorrect grammar.
* **Part of speech tagging**, also called grammatical tagging, is the process of determining the part of speech of a particular word or piece of text based on its use and context. Part of speech identifies ‘make’ as a verb in ‘I can make a paper plane,’ and as a noun in ‘What make of car do you own?’
* **Word sense disambiguation** is the selection of the meaning of a word with multiple meanings  through a process of semantic analysis that determine the word that makes the most sense in the given context. For example, word sense disambiguation helps distinguish the meaning of the verb 'make' in ‘make the grade’ (achieve) vs. ‘make a bet’ (place).
* **Named entity recognition,**or NEM, identifies words or phrases as useful entities. NEM identifies ‘Kentucky’ as a location or ‘Fred’ as a man's name.
* **Co-reference resolution** is the task of identifying if and when two words refer to the same entity. The most common example is determining the person or object to which a certain pronoun refers (e.g., ‘she’ = ‘Mary’),  but it can also involve identifying a metaphor or an idiom in the text  (e.g., an instance in which 'bear' isn't an animal but a large hairy person).
* **Sentiment analysis**attempts to extract subjective qualities—attitudes, emotions, sarcasm, confusion, suspicion—from text.
* **Natural language generation**is sometimes described as the opposite of speech recognition or speech-to-text; it's the task of putting structured information into human language.