



## AI & Machine Learning

[Video Tutorials](#)   [Articles](#)   [Ebooks](#)   [Free Practice Tests](#)   [On-demand Webinars](#)

[Home](#)   [Resources](#)   [AI & Machine Learning](#)   [Deep Learning Tutorial for Beginners](#)   Recurrent Neural Network (RNN) Tutorial: Types, Examples, LSTM and More



# Recurrent Neural Network (RNN) Tutorial: Types, Examples, LSTM and More

Lesson 14 of 18

By Avijeet Biswal

Last updated on Nov 30, 2022

283567

11

## Table of Contents

[What Is a Neural Network?](#)

[What Is a Recurrent Neural Network \(RNN\)?](#)

[Why Recurrent Neural Networks?](#)

[How Does Recurrent Neural Networks Work?](#)

[Feed-Forward Neural Networks vs Recurrent Neural Networks](#)

[View More](#)

Neural Networks is one of the most popular [machine learning algorithms](#) and also outperforms other algorithms in both accuracy and speed. Therefore it becomes critical to have an in-depth understanding of [what a Neural Network is](#), how it is made up and what its reach and limitations are.



## What Is a Neural Network?

A **Neural Network** consists of different layers connected to each other, working on the structure and function of a human brain. It learns from huge volumes of data and uses complex algorithms to train a neural net.

Here is an example of how neural networks can identify a dog's breed based on their features.

- The image pixels of two different breeds of dogs are fed to the input layer of the neural network.
- The image pixels are then processed in the hidden layers for feature extraction.
- The output layer produces the result to identify if it's a German Shepherd or a Labrador.
- Such networks do not require memorizing the past output.

Several neural networks can help solve different business problems. Let's look at a few of them.

- Feed-Forward Neural Network: Used for general **Regression and Classification** problems.
- Convolutional Neural Network: Used for object detection and image classification.
- Deep Belief Network: Used in healthcare sectors for cancer detection.
- RNN: Used for speech recognition, voice recognition, time series prediction, and **natural language processing**.

Read More: [What is Neural Network: Overview, Applications, and Advantages](#)

## Want To Become an AI Engineer? Look No Further!

AI Engineer Master's Program

EXPLORE PROGRAM

## What Is a Recurrent Neural Network (RNN)?



RNN works on the principle of saving the output of a particular layer and feeding this back to the input in order to predict the output of the layer.

Below is how you can convert a Feed-Forward Neural Network into a Recurrent Neural Network:

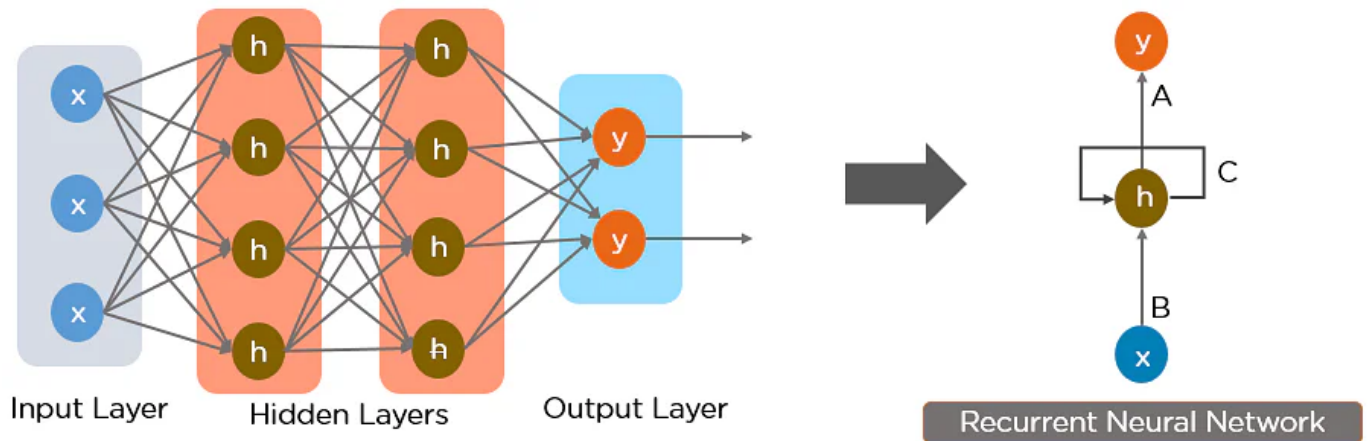
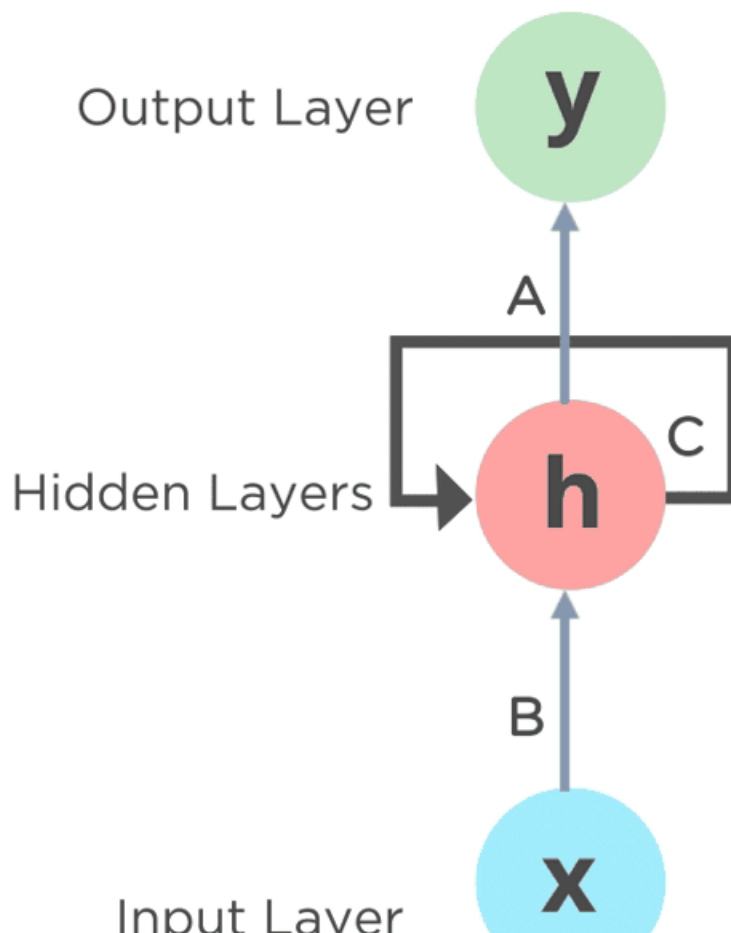


Fig: Simple Recurrent Neural Network

The nodes in different layers of the neural network are compressed to form a single layer of recurrent neural networks. A, B, and C are the parameters of the network.



## A, B and C are the parameters

Fig: Fully connected Recurrent Neural Network

Here, “x” is the input layer, “h” is the hidden layer, and “y” is the output layer. A, B, and C are the network parameters used to improve the output of the model. At any given time t, the current input is a combination of input at x(t) and x(t-1). The output at any given time is fetched back to the network to improve on the output.

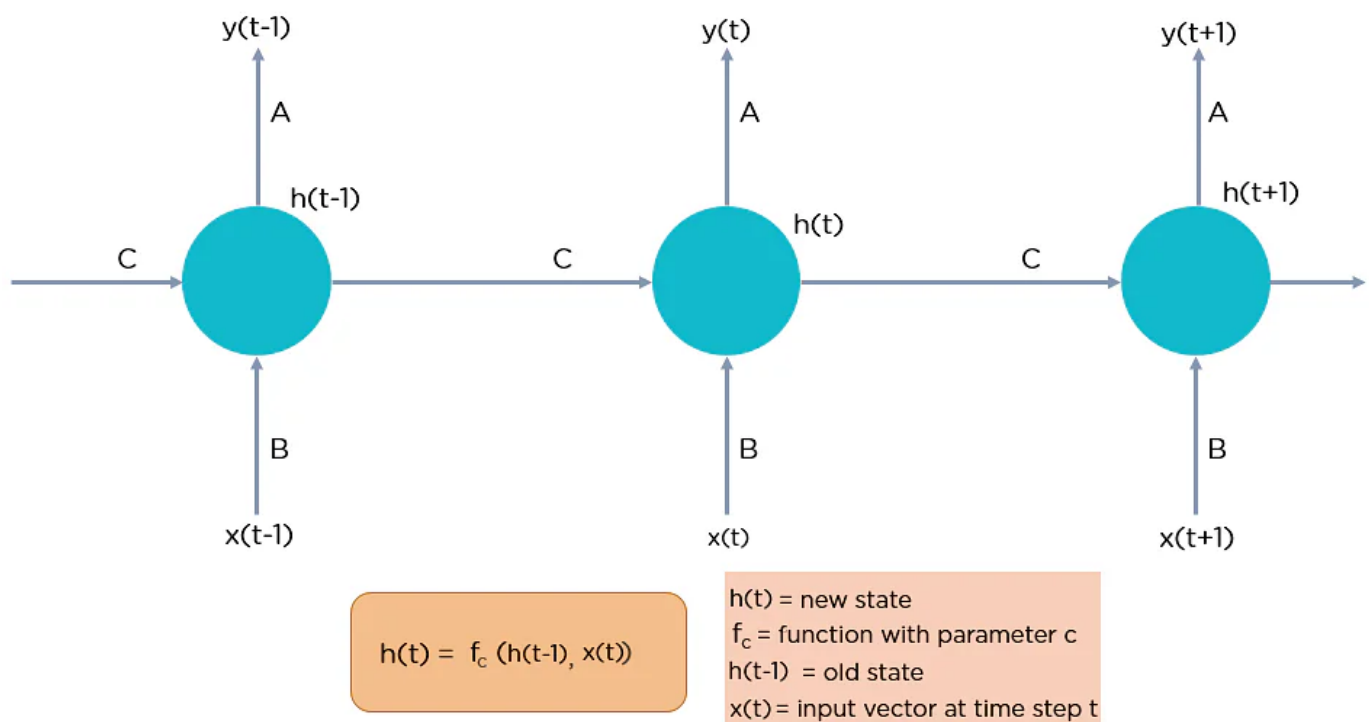


Fig: Fully connected Recurrent Neural Network

Now that you understand what a recurrent neural network is let's look at the different types of recurrent neural networks.

Read More: [An Ultimate Tutorial to Neural Networks](#)

Master deep learning concepts and the TensorFlow open-source framework with the [Deep Learning Training Course](#). Get skilled today!

## Why Recurrent Neural Networks?

RNN were created because there were a few issues in the feed-forward neural network:

- Cannot handle sequential data
- Considers only the current input
- Cannot memorize previous inputs

The solution to these issues is the RNN. An RNN can handle sequential data, accepting the current input data, and previously received inputs. RNNs can memorize previous inputs due to their internal memory.

## Become an AI and ML Expert with Purdue & IBM!

Professional Certificate Program in AI and ML

EXPLORE PROGRAM

### How Does Recurrent Neural Networks Work?

In Recurrent Neural networks, the information cycles through a loop to the middle hidden layer.

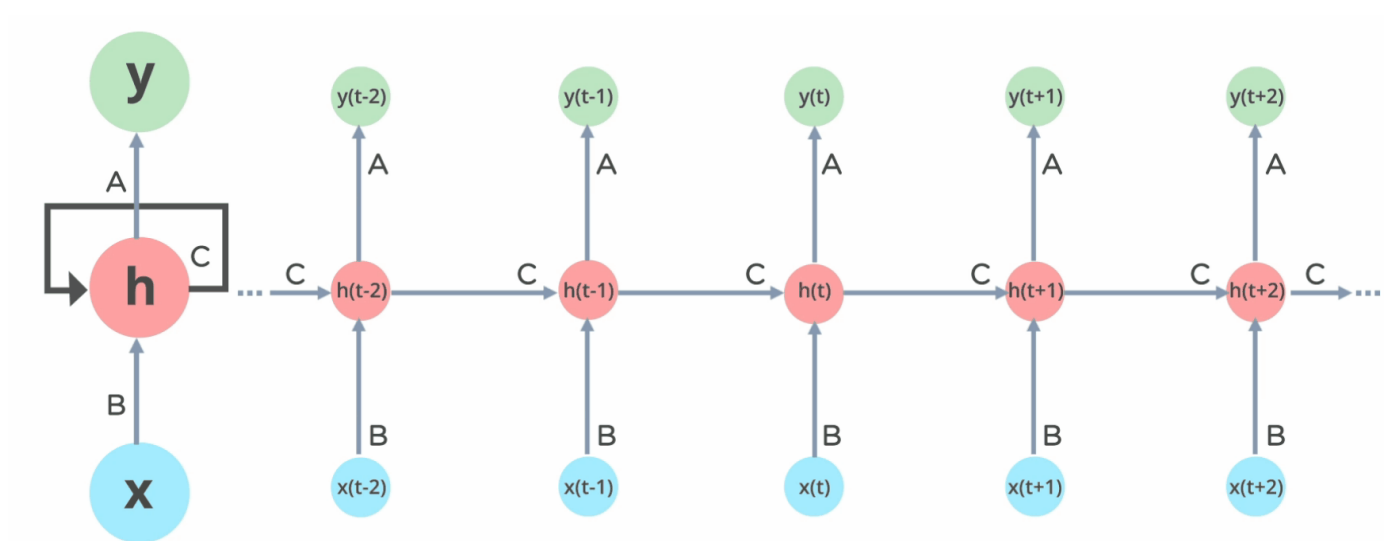


Fig: Working of Recurrent Neural Network

The input layer 'x' takes in the input to the neural network and processes it and passes it onto the middle layer.

The middle layer 'h' can consist of multiple hidden layers, each with its own activation functions and weights and biases. If you have a neural network where the various parameters of different hidden layers are not affected by the previous layer, ie: the neural network does not have memory, then you can use a recurrent neural network.

The Recurrent Neural Network will standardize the different activation functions and weights and biases so that each hidden layer has the same parameters. Then, instead of creating multiple hidden layers, it will create one and loop over it as many times as required.

## Feed-Forward Neural Networks vs Recurrent Neural Networks

A feed-forward neural network allows information to flow only in the forward direction, from the input nodes, through the hidden layers, and to the output nodes. There are no cycles or loops in the network.

Below is how a simplified presentation of a feed-forward neural network looks like:

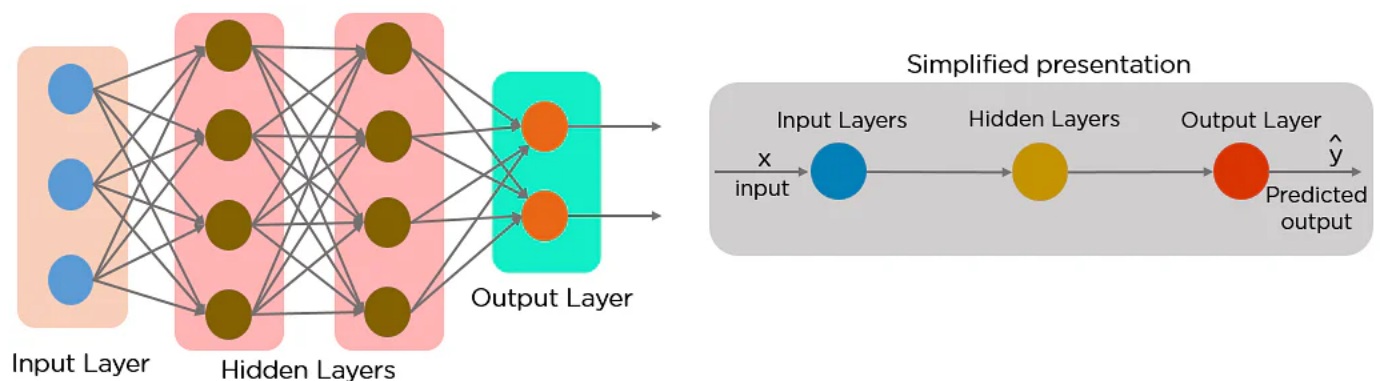


Fig: Feed-forward Neural Network

In a feed-forward neural network, the decisions are based on the current input. It doesn't memorize the past data, and there's no future scope. Feed-forward neural networks are used in general regression and classification problems.

**FREE Machine Learning Certification Course**

To become a Machine Learning Engineer

<https://www.simplilearn.com/tutorials/deep-learning-tutorial/rnn>

[EXPLORE COURSE](#)

## Applications of Recurrent Neural Networks

### Image Captioning

RNNs are used to caption an image by analyzing the activities present.



"A Dog catching a ball in mid air"

### Time Series Prediction

Any time series problem, like predicting the prices of stocks in a particular month, can be solved using an RNN.

### Natural Language Processing

Text mining and Sentiment analysis can be carried out using an RNN for Natural Language Processing (NLP).



When it rains, look for rainbows.  
When it's dark, look for stars.

Positive Sentiment



## Natural Language Processing

### Machine Translation

Given an input in one language, RNNs can be used to translate the input into different languages as output.



Here the person is speaking in English and it is getting translated into Chinese, Italian, French, German and Spanish languages

### Machine Translation

### Types of Recurrent Neural Networks

There are four types of Recurrent Neural Networks:

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

#### One to One RNN

This type of neural network is known as the Vanilla Neural Network. It's used for general machine learning problems, which has a single input and a single output.

one to one



Single output

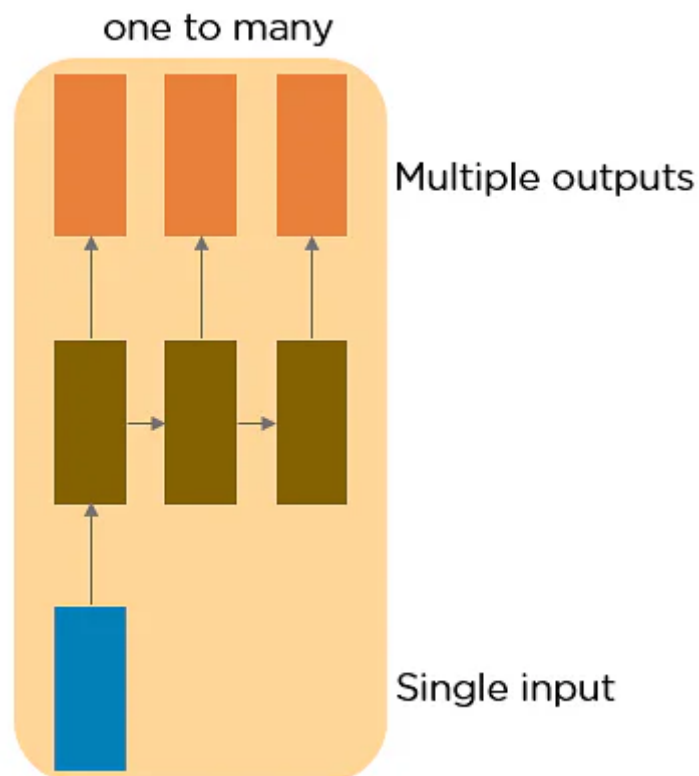
Previous

Next

## Tutorial Playlist

### One to Many RNN

This type of neural network has a single input and multiple outputs. An example of this is the image caption.



Free Webinar | December 20, Tuesday | 9 PM IST

**REGISTER NOW**

## Many to One RNN

This RNN takes a sequence of inputs and generates a single output. Sentiment analysis is a good example of this kind of network where a given sentence can be classified as expressing positive or negative sentiments.



## Many to Many RNN

This RNN takes a sequence of inputs and generates a sequence of outputs. Machine translation is one of the examples.



## Two Issues of Standard RNNs

### 1. Vanishing Gradient Problem

Recurrent Neural Networks enable you to model time-dependent and sequential data problems, such as stock market prediction, machine translation, and text generation. You will find, however, RNN is hard to train because of the gradient problem.

RNNs suffer from the problem of vanishing gradients. The gradients carry information used in the RNN, and when the gradient becomes too small, the parameter updates become insignificant. This makes the learning of long data sequences difficult.



### 2. Exploding Gradient Problem

While training a neural network, if the slope tends to grow exponentially instead of decaying, this is called an Exploding Gradient. This problem arises when large error gradients accumulate, resulting in very large updates to the neural network model weights during the training process.

Long training time, poor performance, and bad accuracy are the major issues in gradient problems.

## Gradient Problem Solutions



Now, let's discuss the most popular and efficient way to deal with gradient problems, i.e., Long Short-Term Memory Network (LSTMs).

First, let's understand Long-Term Dependencies.

Suppose you want to predict the last word in the text: "The clouds are in the \_\_\_\_."

The most obvious answer to this is the "sky." We do not need any further context to predict the last word in the above sentence.

Consider this sentence: "I have been staying in Spain for the last 10 years...I can speak fluent \_\_\_\_."

The word you predict will depend on the previous few words in context. Here, you need the context of Spain to predict the last word in the text, and the most suitable answer to this sentence is "Spanish." The gap between the relevant information and the point where it's needed may have become very large. LSTMs help you solve this problem.

**Join The Fastest Growing Tech Industry Today!**

Professional Certificate Program in AI and ML

**EXPLORE PROGRAM**

## Backpropagation Through Time

Backpropagation through time is when we apply a Backpropagation algorithm to a Recurrent Neural network that has time series data as its input.

In a typical RNN, one input is fed into the network at a time, and a single output is obtained. But in backpropagation, you use the current as well as the previous inputs as input. This is called a timestep and one timestep will consist of many time series data points entering the RNN simultaneously.

Once the neural network has trained on a timeset and given you an output, that output is used to calculate and accumulate the errors. After this, the network is rolled back up and weights are recalculated and updated keeping the errors in mind.

## Long Short-Term Memory Networks

LSTMs are a special kind of RNN — capable of learning long-term dependencies by remembering information for long periods is the default behavior.

All RNN are in the form of a chain of repeating modules of a neural network. In standard RNNs, this repeating module will have a very simple structure, such as a single tanh layer.



Fig: Long Short Term Memory Networks

LSTMs also have a chain-like structure, but the repeating module is a bit different structure. Instead of having a single neural network layer, four interacting layers are communicating extraordinarily.



## Workings of LSTMs in RNN

LSTMs work in a 3-step process.

### Step 1: Decide How Much Past Data It Should Remember

The first step in the LSTM is to decide which information should be omitted from the cell in that particular time step. The sigmoid function determines this. It looks at the previous state ( $h_{t-1}$ ) along with the current input  $x_t$  and computes the function.

Consider the following two sentences:

Let the output of  $h(t-1)$  be "Alice is good in Physics. John, on the other hand, is good at Chemistry."

Let the current input at  $x(t)$  be "John plays football well. He told me yesterday over the phone that he had served as the captain of his college football team."

The forget gate realizes there might be a change in context after encountering the first full stop. It compares with the current input sentence at  $x(t)$ . The next sentence talks about John, so the information on Alice is deleted. The position of the subject is vacated and assigned to John.


## Free Deep Learning for Beginners Course

Master the Basics of Deep Learning

**ENROLL NOW**

### Step 2: Decide How Much This Unit Adds to the Current State

In the second layer, there are two parts. One is the sigmoid function, and the other is the tanh function. In the sigmoid function, it decides which values to let through (0 or 1). tanh function gives weightage to the values which are passed, deciding their level of importance (-1 to 1).




With the current input at  $x(t)$ , the input gate analyzes the important information — John plays football, and the fact that he was the captain of his college team is important.

“He told me yesterday over the phone” is less important; hence it's forgotten. This process of adding some new information can be done via the input gate.

### Step 3: Decide What Part of the Current Cell State Makes It to the Output

The third step is to decide what the output will be. First, we run a sigmoid layer, which decides what parts of the cell state make it to the output. Then, we put the cell state through tanh to push the values to be between -1 and 1 and multiply it by the output of the sigmoid gate.



Let's consider this example to predict the next word in the sentence: “John played tremendously well against the opponent and won for his team. For his contributions, brave \_\_\_\_ was awarded player of the match.”


There could be many choices for the empty space. The current input brave is an adjective, and adjectives describe a noun. So, “John” could be the best output after brave.

## LSTM Use Case


Now that you understand how LSTMs work, let's do a practical implementation to predict the prices of stocks using the “Google stock price” data.

Based on the stock price data between 2012 and 2016, we will predict the stock prices of 2017.

1. Import the required libraries



2. Import the training dataset



### 3. Perform feature scaling to transform the data



### 4. Create a data structure with 60-time steps and 1 output



### 5. Import Keras library and its packages



### 6. Initialize the RNN



### 7. Add the LSTM layers and some dropout regularization.



### 8. Add the output layer.



### 9. Compile the RNN



### 10. Fit the RNN to the training set



### 11. Load the stock price test data for 2017



### 12. Get the predicted stock price for 2017



### 13. Visualize the results of predicted and real stock price



### 15. visualize the results of predicted and real stock price



Looking forward to a successful career in AI and Machine learning. Enrol in our [AI and ML PG Program](#) in collaboration with Purdue University now.

## Next Step to Success

You can also enroll in the [AI and Machine Learning Course](#) with Purdue University and in collaboration with IBM, and transform yourself into an expert in deep learning techniques using [TensorFlow](#), the open-source software library designed to conduct machine learning and deep neural network research. This program in AI and Machine Learning covers [Python](#), Machine Learning, Natural Language Processing, Speech Recognition, Advanced Deep Learning, Computer Vision, and [Reinforcement Learning](#). It will prepare you for one of the world's most exciting technology frontiers.

Have any questions for us? Leave them in the comments section of this tutorial. Our experts will get back to you on the same, as soon as possible.

## About the Author

 [Avijeet Biswal](#)

Avijeet is a Senior Research Analyst at Simplilearn. Passionate about Data Analytics, Machine Learning, and Deep Learning, Avijeet is also interested in politics, cricket, and football.

[View More](#)

## Recommended Programs

Professional Certificate Program in AI and Machine



Learning

4537 Learners

Lifetime  
Access\*



Artificial Intelligence Engineer

24326 Learners

Lifetime  
Access\*



Deep Learning with Keras and TensorFlow

29818 Learners

\*Lifetime access to high-quality, self-paced e-learning content.

[Explore Category](#)

Recommended Resources



Discover the  
Differences  
Between AI vs....



Machine Learning  
Career Guide: A  
compl...

