



# 7 Types of Classification Algorithms in Machine Learning

Classification Algorithms Machine Learning -Explore how classification algorithms work and the types of classification algorithms with their pros and cons.

Last Updated: 27 Sep 2022

[GET ACCESS TO ALL MACHINE LEARNING PROJECTS](#)[VIEW ALL MACHINE LEARNING PROJECTS](#)

*This blog will help you master the fundamentals of classification machine learning algorithms with their pros and cons. You will also explore some exciting machine learning project ideas that implement different types of classification algorithms. So, without much ado, let's dive in.*

Imagine that the pandemic is over and today is a weekday. All the schools, colleges, and offices are open, and you should reach your institution by 8 A.M. You set the alarm last night and managed to wake up at 6 A.M. You have taken your bath and are now nicely dressed in your clothes. You now approach the wall hook to grab your belt, but alas! It's not there. You start panicking and are searching for it here and there. Finally, you call out for your mother, and after 10 minutes of searching, she finds it.



And, now, on the one hand, you are happy that you have found your belt, but on the other hand, you are worried about reaching your institution late. So, on your way to the destination, you start wishing for a magical wardrobe where you could throw your stuff, and it could automatically *classify* things and neatly place them. We all have been through this.



The image shows the homepage of Project Pro's Data Science Projects section. The background is teal with a hexagonal pattern. At the top left is the Project Pro logo. The main title "DATA SCIENCE PROJECTS" is in large white capital letters. To the right is a 3D illustration of a person looking at a large screen displaying data visualizations like line graphs and pie charts. A magnifying glass icon is also present.

## Tensorflow Transfer Learning Model for Image Classification

Downloadable solution code | Explanatory videos | Tech Support

[Start Project](#)

But, if you think that such a wardrobe can't be a reality, let me tell you that's not true. With Artificial Intelligence, it has become possible to build what we like to call 'smart wardrobes.' These will allow you to access things from your wardrobe just by a single tap on your phone. Of course, you might have to explain a few characteristics of your clothing to it initially, for example, its color, size, and type. But, as the wardrobe gradually understands your dress, you will be easily able to utilize it.

### Table of Contents for Classification Machine Learning

- [Types of Classification Tasks in Machine Learning](#)
  - [Binary Classification](#)
  - [Multi-Class Classification](#)
  - [Multi-Label Classification](#)
  - [Imbalanced Classification](#)
- [How Do Classification Algorithms Work?](#)
  - [Classification Algorithms in Machine Learning - Data Preprocessing](#)
  - [Classification Algorithms in Machine Learning-Creating Testing and Training Dataset](#)
  - [Classification Algorithms in Machine Learning -Selecting the Model](#)
- [Types of Classification Algorithms in Machine Learning](#)
  1. [Naive Bayes Classifier](#)
  2. [Logistic Regression](#)
  3. [Decision Tree Classification Algorithm](#)
  4. [Random Forests Classification Algorithm](#)
  5. [Support Vector Machines \(SVMs\)](#)
  6. [K-Nearest Neighbour Classification Algorithm](#)
  7. [K-Means Clustering Classification Algorithm](#)
- [How do you decide which classification algorithm to choose for a given business problem?](#)
- [Classification Algorithms Machine Learning Project Ideas](#)

We expect the wardrobe to perform **classification**, grouping things having similar characteristics together. And there are quite a few classification [machine learning](#) algorithms that can make that happen. We will look through all the different types of classification algorithms in great detail but first, let us begin exploring different types of classification tasks.



## Types of Classification Tasks in Machine Learning

In Machine Learning, most classification problems require predicting a categorical output variable called *target*, based on one or more input variables called *features*. The idea is to fit a statistical model that relates a set of features to its respective target variable to use this model to predict the output for future input observations. However, it is essential to keep in mind that predicting a single output variable will not always be the case. There are many other possible causes, and let us inspect them one by one.

There are mainly four types of classification tasks that one may come across, these are:

1. Binary Classification
2. Multi-Class Classification
3. Multi-Label Classification
4. Imbalanced Classification

**Ace Your Next Job Interview with Mock Interviews from Experts to Improve Your Skills and Boost Confidence!**

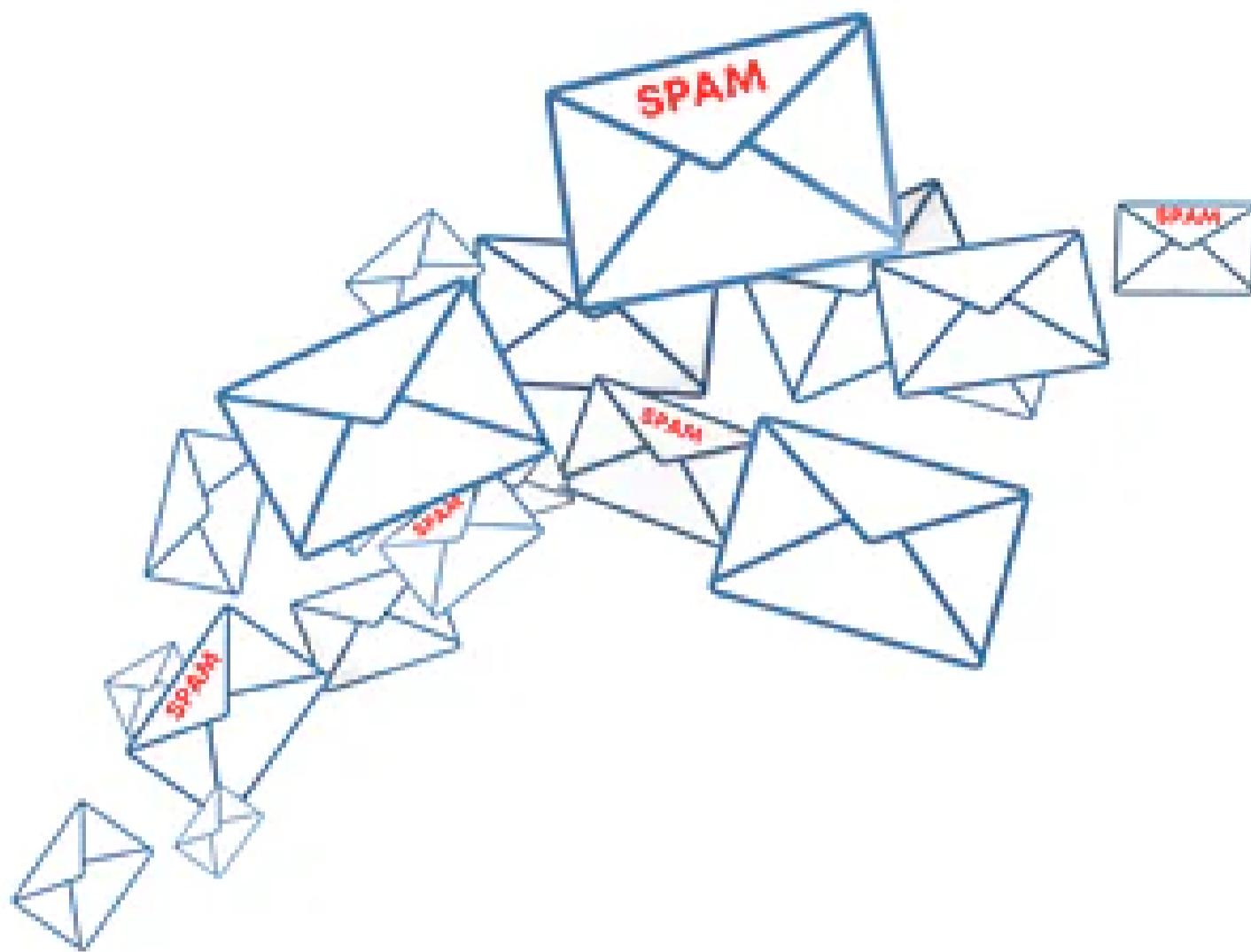
[Get Started](#)

## Binary Classification Machine Learning

This type of classification involves separating the dataset into two categories. It means that the output variable can only take two values.

### Binary Classification Machine Learning Example

The task of labeling an e-mail as "spam" or "not spam." The input variable here will be the content of the e-mail that we are trying to classify. The output variable is represented by 0 for "not spam" and 1 for "spam."



## Multi-Class Classification Machine Learning

In multi-class classification, the output variable can have more than two possible values.

### Multi-Class Classification Machine Learning Example

Identifying the flower type in the case of Iris Dataset where we have four input variables: petal length, sepal length, petal width, sepal width, and three possible labels of flowers: Iris Setosa, Iris Versicolor, and Iris Virginica.

**Iris Versicolor****Iris Virginica****Iris Setosa**

Image Source: Wikipedia Commons

## New Projects

**Learn Performance Optimization Techniques in Spark-Part 2**[View Project](#)**PyCaret Project to Build and Deploy an ML App using Streamlit**[View Project](#)**Build a Data Pipeline in AWS using NiFi, Spark, and ELK Stack**[View Project](#)**Learn to Models using MLLib**[View all New Projects](#)

## Multi-Label Classification Machine Learning

This is an extraordinary type of classification task with multiple output variables for each instance from the dataset. That is, one instance can have multiple labels.

### Multi-Label Classification Machine Learning Example

In [Image classification](#), a single image may contain more than one object, which can be labeled by the algorithm, like bus, car, person, etc.



## Imbalanced Classification

Imbalanced classification refers to classification problems where the instances of the dataset have a biased or skewed distribution. In other words, one class of input variables has a higher frequency than the others.

**Imbalanced Classification Example:** Detecting fraudulent transactions through a credit card in a transaction dataset. Usually, such transactions are remarkably less in number, and this would thus make it difficult for the machine to learn such transactions.



Get Closer To Your Dream of Becoming a Data Scientist with 70+ Solved [End-to-End ML Projects](#)

## How Do Classification Algorithms Work?

To solve classification problems, we use mathematical models that are known as machine learning classification algorithms. Their task is to find how the target variables are related to input features  $x_i$  to the output values,  $y_i$ . In mathematical terms, estimating a function,  $f(x_i)$ , predicts the value of the output variable by taking the associated features as an input. We can write this as,

$$y'_i = f(x_i)$$

where  $y'_i$  represents the predicted response for the output variable.

Now that we understand the task at hand, we will now move forward towards different steps that explain how classification [algorithms in machine learning](#) work.

## Classification Algorithms in Machine Learning - Data Preprocessing

Before we apply any statistical algorithm to our dataset, we must thoroughly understand the input variables and output variables. In classification problems, the target is always qualitative, but sometimes, even the input values can also be categorical, for example, the gender of customers in the famous Mall Customer Dataset. And as classification algorithms are mathematically derived, one must convert all their variables into numerical values. The first step in the working of a classification algorithm is to ensure that the variables, whether input or output, have been encoded correctly.

## Classification Algorithms in Machine Learning-Creating Testing and Training Dataset

After processing the dataset, the next step is to divide the dataset into two parts: the testing dataset and the training dataset. This step allows using the training dataset to make our machine learn the pattern between input and output values. On the other hand, a testing dataset tests the model's accuracy that we will try to fit into our dataset.

## Classification Algorithms in Machine Learning -Selecting the Model

Once we split the dataset into training and testing, the next task is to select the model that best fits our problem. For that, we need to be aware of the popular classification algorithms. So, let us dive into the pool of different types of classification algorithms and explore our options.

## Types of Classification Algorithms in Machine Learning

There are various [classification machine learning](#) algorithms in data mining used by [Data Scientists](#) every day to gain a deeper understanding of their dataset. In this section, we will explore the popular ones in great detail. You are likely to feel like a superhero by understanding the pros and cons of classification algorithms.



Here is a list of different types of classification machine learning algorithms that you will learn about:

1. Naive Bayes Classifier
2. Logistic Regression
3. [Decision Tree](#)
4. Random Forests
5. Support Vector Machines
6. K-Nearest Neighbour
7. K-Means Clustering

## Explore Categories

[Data Science Projects in Python](#)[Data Science Projects in R](#)[Deep Learning Projects](#)[Neural Network Projects](#)[Tensorflow Projects](#)[H2O R Projects](#)[IoT Projects](#)[Keras Deep Learning Projects](#)[NLP Projects](#)[Pytorch](#)

[Data Science Projects in Banking and Finance](#)[Data Science Projects in Retail & Ecommerce](#)[Data Science Projects in Entertainment & Media](#)[Data Science Projects in Telecommunications](#)

## 1. Naive Bayes Classifier

Naive Bayes classifier, is one of the simplest and most [effective classification machine learning](#) algorithms. Its basis is Bayes' theorem which describes how the probability of an event is evaluated based on prior knowledge of conditions that might be related to the event. Mathematically, this theorem states-

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

Where  $P(Y|X)$  is the probability of an event Y, given that even X has already occurred.

$P(X)$  is the probability of event X,

$P(Y)$  is the probability of event Y,

$P(X|Y)$  is the likelihood of event X given a fixed value of Y.

If X represents a feature variable and Y represents a target variable, then the Bayes Classifier will assign that label to the feature that produces the highest probability. For simplicity, consider a two-class problem where the feature variable can have only two possible values,  $Y=1$  or  $Y=0$ . Then, the Bayes Classifier will predict class 1 if  $\Pr(Y=1|X=x_0) > 0.5$ , and class two otherwise.

In cases of more than one feature, we can use the following formula for evaluating the probability,

$$P(Y|X_1 = x_1, X_2 = x_2) = \frac{P(X_1 = x_1, X_2 = x_2|Y)P(Y)}{P(X_1 = x_1, X_2 = x_2)}$$

where we have assumed that the two features  $X_1$  and  $X_2$ , are independent of each other. In fact, because of this assumption, the word 'Naive' is attached to Bayes' classifier.

### Naive Bayes Classification Algorithms Example

Consider the following dataset where a sportsperson plays or not was observed along with the weather conditions.

Humidity	Wind	Whether the sportsperson will go out to play?
High	Weak	No
High	Strong	No
High	Weak	Yes
High	Weak	Yes
Normal	Weak	Yes
Normal	Strong	No
Normal	Strong	Yes
High	Weak	No
Normal	Weak	Yes
Normal	Weak	Yes
Normal	Strong	Yes
High	Strong	Yes
Normal	Weak	Yes
High	Strong	No

Data Source: Mitchell, T. M. (1997). Machine learning.

Suppose we now have to predict whether the person will play or not, given that humidity is 'High' and the wind is 'Strong.' Then, using the Bayes' classifier, we can compute the probability as follows:

$$P(Y = 'Yes' | X_1 = 'High', X_2 = 'Strong') = \frac{P(X_1 = 'High', X_2 = 'Strong' | Y = 'Yes') P(Y = 'Yes')}{P(X_1 = 'High', X_2 = 'Strong')}$$

$$= \frac{\frac{1}{9} \times \frac{9}{14}}{\frac{3}{14}}$$

$$= \frac{1}{3}$$

$$P(Y = 'No' | X_1 = 'High', X_2 = 'Strong') = \frac{P(X_1 = 'High', X_2 = 'Strong' | Y = 'No') P(Y = 'No')}{P(X_1 = 'High', X_2 = 'Strong')}$$

$$= \frac{\frac{2}{5} \times \frac{5}{14}}{\frac{3}{14}}$$

$$= \frac{2}{3}$$

## Advantages of Naive Bayes Classification Algorithm

1. It is simple, and its implementation is straightforward.
2. The time required by the machine to learn the pattern using this classifier is less.
3. It performs well in the case where the input variables have categorical values.
4. It gives good results for complex real-world problems.
5. It performs well in the case of multi-class classification.

## Disadvantages of Naive Bayes Classification Algorithm

1. It assumes independence among feature variables which may not always be the case.
2. We often refer to it as a bad estimator, and hence the probabilities are not always of great significance.
3. If, during the training time, the model was not aware of any of the categorical variables and that variable is passed during testing, the model assigns 0 (zero) likelihood and thus substitutes zero probability referred to as 'zero frequency.' One can avoid this situation by using smoothing procedures such as Laplace estimation.

## Applications of Naive Bayes Classification Algorithm

1. **Spam Classification:** Identifying whether an e-mail is a spam or not based on the content of the e-mail
2. **Live Prediction System:** This model is relatively fast and thus predicts the target variable in real-time.
3. **Sentiment Analysis:** Recognising feedback on a product and classifying it as 'positive' or 'negative.'
4. **Multi-Class Prediction:** Naive Bayes works well for multi-class classification machine learning problems.

Get FREE Access to [Machine Learning Example Codes](#) for Data Cleaning, Data Munging, and Data Visualization

## 2. Logistic Regression

This algorithm is similar to Bayes' classifier as it also predicts the probability that Y is associated with an input variable, X. It uses the logistic function,

$$P(X) = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}}$$

and fits the parameters  $\beta_0$  and  $\beta_1$  using the maximum likelihood technique. This technique involves maximizing the likelihood function given by

$$l(\beta_0, \beta_1) = \prod_{l:y_i=1} p(x_i) \prod_{l':y_{i'}=1} (1 - p(x_i))$$

After evaluating the two parameters, one can easily use the logistic function to predict the target variable probability  $p(x_i)$  for a given input  $x_i$ .

In the case of more than one feature variables ( $X_1, X_2, \dots, X_p$ ), the formula can be generalized as

$$P(X) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}$$

## Advantages of Logistic Regression Classification Machine Learning Algorithm

1. It's a simple model, so it takes very little time for training.
2. It can handle a large number of features.

## Disadvantages of Logistic Regression Classification Algorithm

1. Although it has the word regression in its name, we can only use it for classification problems because of its range which always lies between 0 and 1.
2. It can only be used for binary classification problems and has a poor response for multi-class classification problems

## Applications of Logistic Regression Classification Algorithm

1. **Credit Scoring:** To predict the creditworthiness (ability to pay back borrowed loan) of an individual based upon some features like annual income, account balance, etc.
2. **Predicting User Behaviour:** Many websites use logistic regression to predict user behavior and guide them towards clicking on links that might interest them.
3. **Discrete Choice Analysis:** Logistic regression is an excellent choice for predicting the categorical preferences of people. Examples for this could be which car to buy, which school or college to attend, etc., based on the attributes of people and the diverse options available to them.

### Recommended Reading:

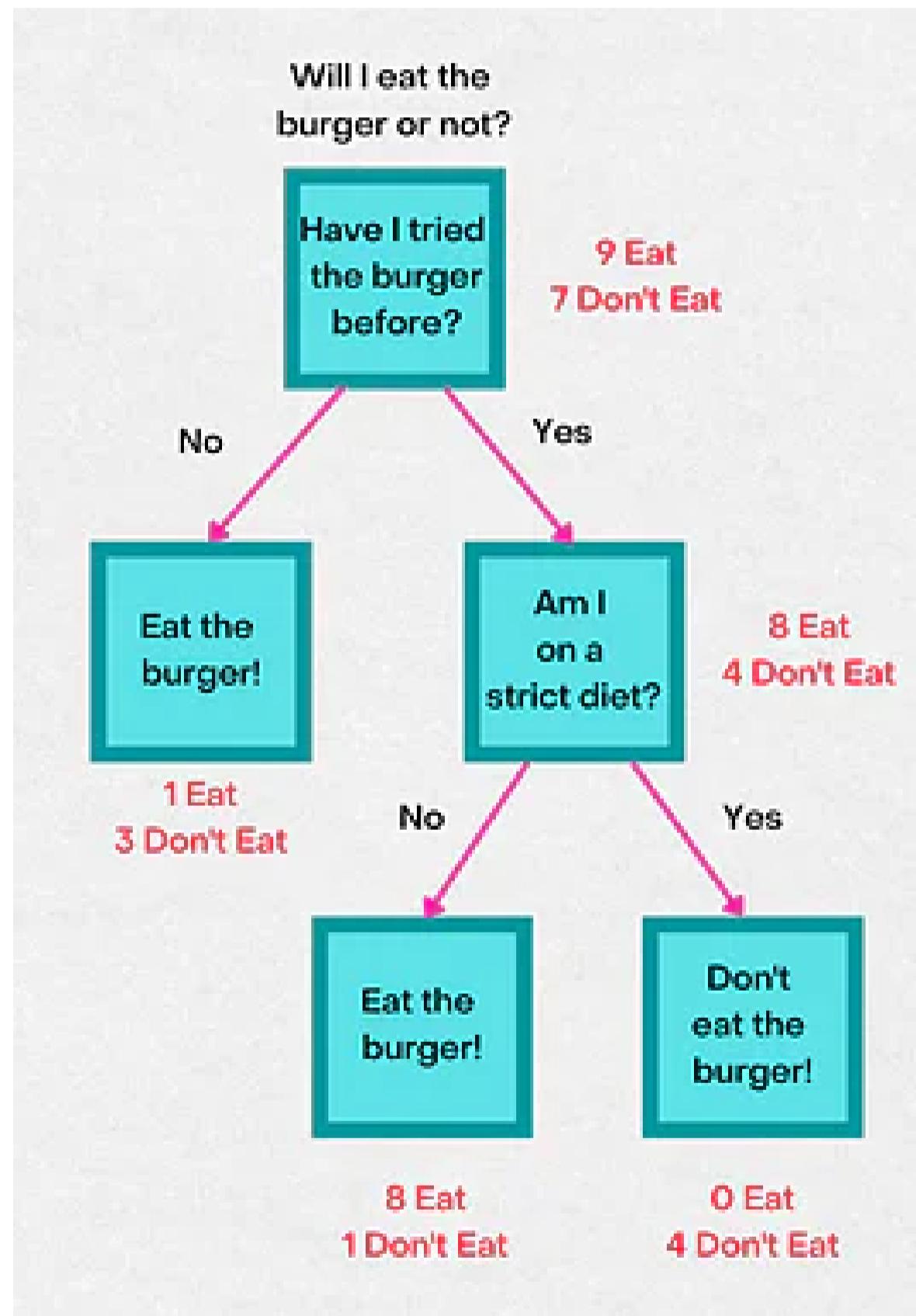
- [The A-Z Guide to Gradient Descent Algorithm and Its Variants](#)
- [Ensemble Learning](#)
- [Types of Neural Networks](#)

- [5 Different Types of Neural Networks](#)
- [15 Time Series Projects Ideas for Beginners to Practice 2021](#)
- [Exploratory Data Analysis in Python-Stop, Drop and Explore](#)
- [How to Become an Artificial Intelligence Engineer in 2021](#)
- [8 Feature Engineering Techniques for Machine Learning](#)
- [AI Engineer Salary - The Ultimate Guide for 2021](#)
- [Logistic Regression vs Linear Regression in Machine Learning](#)
- [Correlation vs. Covariance](#)

### 3. Decision Tree Classification Algorithm

This classification machine learning algorithm involves dividing a dataset into segments based on certain feature variables from the dataset. The threshold values for these divisions are usually the mean or mode of the respective feature variable (if they are numerical). As a tree can represent the set of splitting rules used to segment the dataset, this algorithm is known as a decision tree.

Look at the example below to understand it better.



The text in red represents how the dataset has been split into segments based on the output variable. The outcome is the one that has the highest proportion of a class.

Now, the question which is quite natural to ask is what criteria this algorithm uses to split the data. There are two widely used measures to test the purity of the split (a segment of the dataset is pure if it has data points of only one class).

The first one is the Gini index defined by

$$G = - \sum_{k=1}^k p_{mk} (1 - p_{mk})$$

that measures total variance across the N classes. Another measure is cross-entropy, defined by

$$D = - \sum_{k=1}^k p_{mk} \log(p_{mk})$$

In both equations,  $p_{mk}$  represents the proportion of training variables in the mth segment that belongs to the kth class.

We split the dataset into segments based on that feature, giving rise to the minimum value of entropy or Gini index.

## Get confident to build end-to-end projects.

Access to a curated library of 250+ end-to-end industry projects with solution code, videos and tech support.

[Request a demo](#)

### Advantages of Decision Tree Classification Algorithm

1. This algorithm allows for an uncomplicated representation of data. So, it is easier to interpret and explain it to executives.
2. Decision Trees mimic the way humans make decisions in everyday life.
3. They smoothly handle qualitative target variables.
4. They handle non-linear data effectively.

### Disadvantages of Decision Tree Classification Algorithm

1. They may create complex trees which sometimes become irrelevant.
2. They do not have the same level of prediction accuracy as compared to other algorithms.

### Applications of Decision Tree Classification Algorithm

1. Sentiment Analysis: It is used as a classification algorithm in text mining to determine a customer's sentiment towards a product.
2. Product Selection: Companies can use decision trees to realize which product will give them higher profits on launching.

## 4. Random Forests Classification Algorithm

A forest consists of a large number of trees. Similarly, a random forest involves processing many decision trees. Each tree predicts a value for the probability of target variables. We then average the probabilities to produce the final output.

We evaluate each tree as follows:

1. First samples of the dataset are created by selecting data points with replacement.
2. Next, we do not use all input variables to create decision trees. We use only a subset of the available ones.
3. Each tree is allowed to grow to the most considerable length possible, and no pruning is involved.

### Advantages of Random Forest Classification Algorithm

1. It is efficient when it comes to large datasets.
2. It allows estimating the significance of input variables in classification.
3. It is more accurate than decision trees.

### Disadvantages of Random Forest Classification Algorithm

1. It is more complex when it comes to implementation and thus takes more time to evaluate.

### Applications of Random Forest Classification Algorithm

1. **Credit Card Default:** Credit card companies use random forests to predict whether the cardholder will default on their debt or not.
2. **Stock Market Prediction:** Stock investors use it to indicate a particular stock's trends and analyze loss and profit from it.
3. **Product Recommendation:** One can use it to recommend products to a user based on their preferences.

Get More Practice, More [Data Science and Machine Learning Projects](#), and More guidance. Fast-Track Your Career Transition with ProjectPro

## 5. Support Vector Machines (SVMs)

This algorithm utilizes [support vector](#) classifiers with an exciting change that makes it suitable for evaluating a non-linear decision boundary. And that becomes possible by enlarging the feature variable space using special functions called *kernels*. The decision boundary that this algorithm considers allows labeling the feature variable to a target variable. The mathematical function that it uses for evaluating the boundaries is given by

$$f(x) = \beta_0 - \sum_{i \in s} \alpha_i K(x, x_i)$$

where  $K$  represents the kernel function, and  $\alpha_i$  and  $\beta_0$  are training parameters.

### Advantages of SVM Classification Algorithm

1. It makes training the dataset easy.
2. It performs well when the data is high-dimensional.

### Disadvantages of SVM Classification Algorithm

1. It doesn't perform well when the data has noisy elements.
2. It is sensitive to kernel functions, so they have to be chosen wisely.

### Applications of SVM Classification Algorithm

1. **Face Detection:** It is used to read through images (an array of pixel numbers) and identify whether it contains a face or not based on usual human features.
2. **Image Classification:** [SVM](#) is one of the image classification algorithms used to classify images based on their characteristics.
3. **Handwritten Character Recognition:** We can use it to identify handwritten characters.

**Time for a tea break!**



Dear reader, please feel free to take a short break now to celebrate that you have learned so much about classification algorithms. FYI, till now, the algorithms that we have discussed are all instances of supervised classification algorithms. These are the algorithms that have pre-defined target variables for them in the dataset. But, in the real world, this may not always be the case. So, we will now explore unsupervised classification machine learning algorithms where the task at hand would be to learn the pattern among input features and group together similar ones.

## Most Watched Projects

[Build an AWS ETL Data Pipeline in Python on YouTube Data](#)

[View Project](#)

[PySpark Project-Build a Data Pipeline using Kafka and Redshift](#)

[View Project](#)

[Linear Regression Model Project in Python for Beginners Part 1](#)

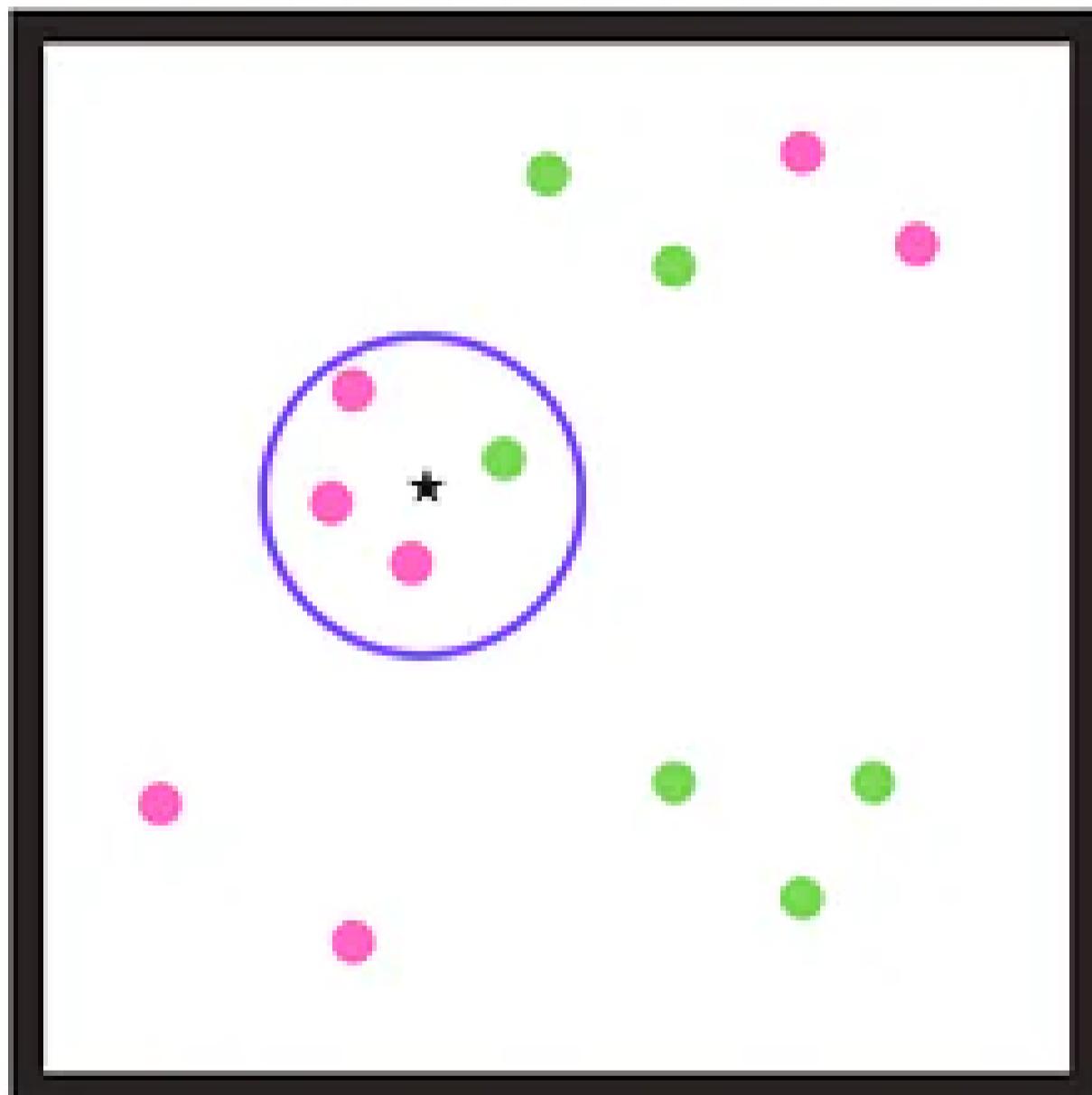
[View Project](#)

[AWS SNS Example Airflow](#)

[View all Most Watched Projects](#)

## 6. K-Nearest Neighbour Classification Algorithm

KNN algorithm works by identifying K nearest neighbors to a given observation point. It then evaluates the proportions of each type of target variable using the K points and then predicts the target variable with the highest ratio. For example, consider the following case where we have to label a target value to point X. Then, if we take four neighbors around it, this model will predict that the point belongs to class with the color pink.



### Advantages of K-Nearest Neighbour Classification Algorithm

1. One can apply it to datasets of any distribution.
2. It is easy to understand and is quite intuitive.

### Disadvantages of K-Nearest Neighbour Classification Algorithm

1. It is easily affected by outliers.
2. It is biased towards a class that has more instances in the dataset.
3. It is sometimes challenging to find the optimal number for K.

## Applications of KNN Classification Algorithm

1. **Detecting Outliers:** As the algorithm is sensitive to outlier instances, it can detect outliers.
2. **Identifying Similar Documents:** To recognize semantically similar documents.

## 7. K-Means Clustering Classification Algorithm

[K-Means](#) Clustering is a clustering algorithm that divides the dataset into K non-overlapping groups. The first step for this algorithm is to specify the expected number of clusters, K. Then, the task is to divide the dataset into K clusters so that within-the-cluster variation is as tiny as possible. The algorithm proceeds as follows:

1. Assign a number from 1 to K randomly to each input variable. These are initial cluster labels for the variables.
2. Repeat the step until the cluster assignments remain unchanged:
3. Evaluate the cluster centroid for each of the K clusters.
4. Assign each input variable set to the cluster whose centroid is closest (here closest can be defined in terms of Euclidean distance)

In conclusion, this classification machine learning algorithm minimizes the sum of squares of deviations between an input point and the respective cluster centroid. The reason for naming it as [K-means clustering](#) is that step 2a) evaluates the mean of the observations belonging to a particular cluster as the cluster centroid.

## Advantages of KNN Classification Algorithm

1. We can apply it to large datasets.
2. It is effortless to implement.
3. It guarantees convergence for locating clusters.

## Disadvantages of KNN Classification Algorithm

1. It has a limitation as one has to provide the value for K initially.
2. It is sensitive to outliers.

## Applications of KNN Classification Algorithm

1. **Add Recommendation:** Companies can identify clusters of customers who share money spending habits and present advertisements that they are more likely to buy.
2. **Identifying crime zones in a city:** Using [K-means clustering](#), we can identify areas more prone to criminal cases.
3. **Document Classification:** To identify documents written on a similar topic.

Now that you are familiar with so many different classification algorithms, it is time to understand which one to use when.

[Access Data Science and Machine Learning Project Code Examples](#)

## How do you decide which classification algorithm to choose for a given business problem?

Below we have a list that will help you understand which classification machine learning algorithms you should use for solving a business problem.

1. **Problem Identification:** The first and foremost thing to do would be to understand the task at hand thoroughly. If it's a supervised classification case, you can use algorithms like Logistic Regression, Random Forest, Decision Tree, etc. On the other hand, if it is an unsupervised classification case, you should go for clustering algorithms.
2. **Size of the dataset:** The size of the dataset is also a parameter that you should consider while selecting an algorithm. As few algorithms are relatively fast, it'll be better to switch to those. If the size of the dataset is small, you can stick to low bias/high variance algorithms like Naive Bayes. In contrast, if the dataset is large, the number of features is high, then you should use high bias/low variance algorithms like [KNN](#), Decision trees, and SVM.
3. **Prediction Accuracy:** The accuracy of a model is a parameter that tests how good a classifier is. It reflects how well the predicted output value matches the correct output value. Of course, higher accuracy is desirable, but one should also check that the model does not overfit.
4. **Training Time:** Sometimes, complex classification machine learning algorithms like SVM and Random Forests may take up a lot of time for computation. Also, higher accuracy and large datasets anyway require more time to learn the pattern. Simple algorithms like Logistic Regression are easier to implement and save time.
5. **Linearity of the Dataset:** Not always is there a linear relationship between input variables and target variables. It is thus essential to analyze this relationship and choose the algorithm carefully as a few of them are restricted to linear datasets. The best method to

check for linearity is to either fit a linear line or run a logistic regression or SVM and look for residual errors. A higher error suggests the data is non-linear and would require the implementation of complex algorithms.

**6. Number of Features:** Sometimes, the dataset may contain unnecessarily many features, and not all of them will be relevant. One can then use algorithms like SVM, best suited for such cases, or use Principal Component Analysis to figure out which features are significant.

## Classification Algorithms Machine Learning Project Ideas

Now you are entirely ready to explore some [hands-on machine learning projects](#) which implement these algorithms for solving real-world problems.

1. [E-Commerce Product Reviews- Pairwise Ranking and Sentiment Analysis](#): In this ML project, you will understand how to use machine learning algorithms for text classification in Python. You will also realize how one can use a classification problem to recognize the sentiment of a user.
2. [TalkingData Ad tracking Fraud Detection](#): This is an insightful [machine learning project idea](#) that will build your understanding of two classification algorithms, Decision Tree and Logistic Regression. You will learn about building a fraud detection system from scratch that detects whether a click on an advertisement will result in fraud or not.
3. [How to prepare test data for your machine learning project](#): This is another exciting project that will help you explore implementing a multi-class classification solution in the practical world. It will also help you realize how we use machine learning classification algorithms in Natural Language Processing (NLP) related problems.
4. [Predicting Loan Default](#): This ML project will introduce you to the application of Random Forest and Logistic Regression to predict the loan eligibility based on the data entered by him/her.
5. [NLP Projects - Kaggle Quora Question Pairs Solution](#): This [NLP project](#) implements a Random Forest classifier to identify which questions are similar.

Well, we now hope that you've mastered the concepts of classification algorithms and feel like a superhero.

[PREVIOUS](#)

[NEXT](#)

Your Data Skills Need to Get Stronger  
And We Have The Projects Ready

GET ACCESS TO SOLVED PROJECTS

### Start Your First Project

Learn By Doing

Your phone

Your email

START PROJECT

## Trending Project Categories

Machine Learning Projects  
Data Science Projects  
Deep Learning Projects  
Big Data Projects  
Apache Hadoop Projects  
Apache Spark Projects  
[Show more](#)

## Trending Projects

Walmart Sales Forecasting Data Science Project  
BigMart Sales Prediction ML Project  
Music Recommender System Project  
Credit Card Fraud Detection Using Machine Learning  
Resume Parser Python Project for Data Science  
Time Series Forecasting Projects  
[Show more](#)

## Trending Blogs

Machine Learning Projects for Beginners with Source Code  
Data Science Projects for Beginners with Source Code  
Big Data Projects for Beginners with Source Code  
IoT Projects for Beginners with Source Code  
Data Analyst vs Data Scientist  
Data Science Interview Questions and Answers  
[Show more](#)

## Trending Recipes

Search for a Value in Pandas DataFrame  
Pandas Create New Column based on Multiple Condition  
LSTM vs GRU  
Plot ROC Curve in Python  
Python Upload File to Google Drive  
Optimize Logistic Regression Hyper Parameters  
[Show more](#)

## Trending Tutorials

PCA in Machine Learning Tutorial  
PySpark Tutorial  
Hive Commands Tutorial  
MapReduce in Hadoop Tutorial  
Apache Hive Tutorial -Tables  
Linear Regression Tutorial  
[Show more](#)

© 2022 Iconiq Inc.

[About us](#)

[Contact us](#)

[Privacy policy](#)

[User policy](#)

[Write for ProjectPro](#)

**ProjectPro**