

RESEARCH

PROGRAMMING

ARTIFICIAL INTELLIGENCE

INTERVIEWS

OTHER

# Classification Algorithms: A Tomato-Inspired Overview >



Article by Yulia Gavrilova  
August 5th, 2020

7 min read

  37

Remember the No Free Lunch theorem? No, it is not about food (yet). But if you are hungry, get a snack before reading this post – I don't want you drooling all over your keyboard.

I will remind you – no algorithm is optimal over the set of all possible situations. Machine learning algorithms are delicate instruments that you tune based on the problem set, especially in supervised machine learning.

Today, we will see how popular classification algorithms work and help us, for example, to pick out and sort wonderful, juicy tomatoes.

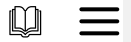
## How classification works >

We predict whether a thing can be referred to a particular class every day. To give an example, classification helps us make decisions when picking tomatoes in a supermarket ("green", "perfect", "rotten"). In machine learning terms, we assign a label of one of the classes to every tomato we hold in our hands.

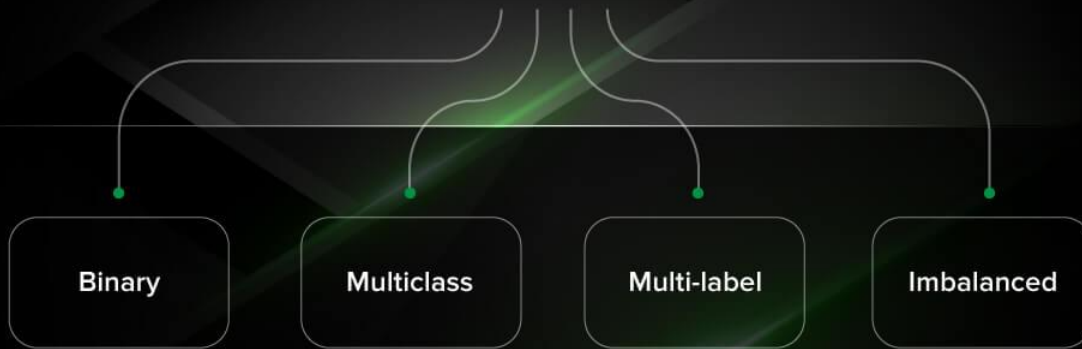
The efficiency of your Tomato Picking Contest (some would call it a classification model) depends on the accuracy of its results. The more often you go to the supermarket yourself (instead of sending your parents or your significant other), the better you will become at picking out tomatoes that are fresh and yummy.

Computers are just the same! For a classification model to learn to predict outcomes accurately, it needs a lot of training examples.

## 4 types of classification >



## 4 types of classification:



### Binary



Binary classification means there are two classes to work with that relate to one another as true and false. Imagine you have a huge lug box in front of you with yellow and red tomatoes. But, your fancy Italian pasta recipe says that you only need the red ones.

What do you do? Obviously, you use **label-encoding** and, in this case, assign 1 to “red” and 0 to “not red”. Sorting tomatoes has never been easier.



## Algorithms for binary classification:

Logistic  
Regression

k-Nearest  
Neighbors

Decision  
Trees

Support Vector  
Machine

Naive Bayes

## Multiclass



What do you see in this photo?





Red beefsteak Tomatoes. Cherry tomatoes. Cocktail tomatoes. Heirloom tomatoes.

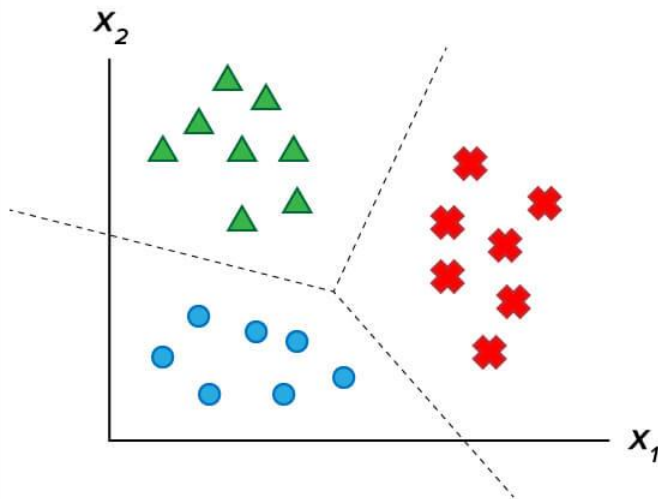


There is no black and white here, “normal” and “abnormal” like in binary classification. We welcome all sorts of wonderful vegetables (or berries) to our table.

What you probably don’t know if you are not a fan of tomatoes cooking is that not all the tomatoes are equally good for the same dish. Red beefsteak tomatoes are perfect for salsa but you do not pickle them. Cherry tomatoes work for salads but not for pasta. So it is important to know *what* you are dealing with.

Multiclass classification helps us to sort all the tomatoes from the box regardless of how many classes there are.

## Multiclass classification algorithms



**k-Nearest Neighbors,  
Decision Trees,  
and Naive Bayes.**

**Random Forest,  
Gradient Boosting,  
Artificial neural networks.**

## Multi-label

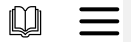


Multi-label classification is applied when one input can belong to more than one class, like a person who is a citizen of two countries.

To work with this type of classification, you need to build a model that can predict multiple outputs.

You need a multi-label classification for object recognition in photos. For example, when you need to identify not only tomatoes but also different other kinds of objects in the same image, such as mushrooms.

identify not only tomatoes but also different other kinds of objects in the same image: apples, zucchinis, onions etc.



**Important note for all tomato lovers:** You cannot just take a binary or multiclass classification algorithm and apply it directly to multi-label classification. But you can use:

You can also try to use a separate algorithm for each class to predict labels for each category.

## Imbalanced



We work with Imbalanced classification when examples in each class are unequally distributed.

Imbalanced classification is used for fraud detection software and medical diagnosis. Finding rare and exquisite biologically grown tomatoes that are accidentally spilled in a large pile of supermarket tomatoes is an example of imbalanced classification offered by Gints, our awesome editor (if you have any other examples, [tweet them to us](#)).

I recommend you visit the fantastic blog of [Machine Learning Mastery](#) where you can read about the different types of classification and study many more machine learning materials.

## Steps to build a classification model



Once you know what kind of classification task you are dealing with, it is time to build a model.

1. Select the classifier. You need to choose one of the ML algorithms that you will apply to your data.
2. Train it. You have to prepare a training data set with labeled results (the more examples, the better).
3. Predict the output. Use the model to get some results.
4. Evaluate the classifier model. It is a good idea to prepare a validation set of data that you have not used in training to check the results.

Let us now take a look at the most widely-used classification algorithms.

## The most popular classification algorithms



Scikit-Learn is one of the top ML libraries for Python. So if you want to build your model, check it out. It provides access to widely-used classifiers.

## Logistic Regression



[Logistic regression](#) is used for binary classification.

This algorithm employs a logistic function to model the probability of an outcome happening. It is most useful when you want to understand how several independent variables affect a single outcome variable.

Example question: Will the precipitation levels and the soil composition lead to tomato's prosperity or untimely death?



Logistic regression has limitations; all predictors should be independent, and there should be no missing values. This algorithm will fail when there is no linear separation of values.

## Naive Bayes



The Naive Bayes algorithm is based on the **Bayes' theorem**. You can apply this algorithm for binary and multiclass classification and classify data based on historical results.

Example task: I need to separate rotten tomatoes from the fresh ones based on their look.

The advantages of Naive Bayes are that these algorithms are fast to build: they do not require an extensive training set and are also fast compared to other methods. However, since the performance of Bayesian algorithms depends on the accuracy of its strong assumptions, the results can potentially turn out very bad.

Using Bayes' theorem, it is possible to tell how the occurrence of an event **impacts the probability** of another event.

### Naive Bayes, Clearly Explained!!!



## k-Nearest Neighbors



kNN stands for "k-nearest neighbor" and is one of the simplest classification algorithms.

The algorithm assigns objects to the class that most of its nearest neighbors in the **multidimensional feature space** belong to. The number k is the number of neighboring objects in the feature space that are compared with the classified object.



are compared with the classified object.



Example: I want to predict the species of the tomato from the species of tomatoes similar to it.



To classify the inputs using k-nearest neighbors, you need to perform a set of actions:

- Calculate the distance to each of the objects in the training sample;
- Select k objects of the training sample, the distance to which is minimal;
- The class of the object to be classified is the class that occurs most frequently among the k-nearest neighbors.

## Decision Tree



Decision trees are probably the most intuitive way to visualize a decision-making process. To predict a class label of input, we start from the root of the tree. You need to divide the possibility space into smaller subsets based on a decision rule that you have for each **node**.

Here is an example:

You keep breaking up the possibility space until you reach the bottom of the tree. Every decision node has two or more branches. The leaves in the model above contain the decision about whether a person is or isn't fit.

Example: You have a basket of different tomatoes and want to choose the correct one to enhance your dish.

## Types of Decision Trees



There are two types of trees. They are based on the nature of the target variable:

- Categorical Variable Decision Tree.
- Continuous Variable Decision Tree.

Therefore, decision trees work quite well with both numerical and categorical data. Another plus of using decision trees is that they require little data preparation.

However, decision trees can become too complicated, which leads to **overfitting**. A significant disadvantage of these algorithms is that small variations in training data make them unstable and lead to entirely new trees.

## Random Forest



Random forest classifiers use several different decision trees on various sub-samples of datasets. The average result is taken as the model's prediction, which improves the predictive accuracy of the model in general and combats overfitting.

Random Forest Algorithm - Random Forest Explained | Random Forest in ...



Consequently, random forests can be used to solve complex machine learning problems without compromising the accuracy of the results. Nonetheless, they demand more time to form a prediction and are more challenging to implement.

Read more about how random forests work in the [Towards Data Science blog](#).

## Support Vector Machine



Support vector machines use a hyperplane in an N-dimensional space to classify the data points. N here is the number of features. It can be, basically, any number, but the bigger it is, the harder it becomes to build a model.

### Support Vector Machines Part 1 (of 3): Main Ideas!!!







One can imagine the hyperplane as a line (for a 2-dimensional space). Once you pass 3-dimensional space, it becomes hard for us to visualize the model.

Data points that fall on different sides of the hyperplane are attributed to different classes.

Example: An automatic system that sorts tomatoes based on their shape, weight, and color.

The hyperplane that we choose directly affects the accuracy of the results. So, we search for the plane that has the maximum distance between data points of both classes.

SVMs show accurate results with minimal computation power when you have a lot of features.

## Summing up




As you can see, machine learning can be as simple as picking up vegetables in the shop. But there are many details to keep in mind if you don't want to mess it up. Stay tuned to our blog, [Twitter](#), and [Medium](#) for more cool reads about machine learning.

TAGGED: [algorithms](#) [machine learning](#)

Share:



 37 upvotes

Suggest your topic for our next post...

Submit



More from Serokell





## How to Choose a Machine Learning Technique

Need to build an ML model but don't know where to start? In this post, we will help you pick the correct machine learning algorithms for your particular use case.

July 8th, 2020 | 8 min read



## Top Machine Learning Career Paths in 2021

Learn about awesome artificial intelligence and machine learning jobs that you can pursue in 2021.

May 10th, 2021 | 8 min read



## Machine Learning Algorithm Classification for Beginners

Read about the most commonly used machine learning algorithms and how they are categorized.

June 11th, 2020 | 14 min read



Biotech

Managed IT services

ML consulting

Fintech

How we work

Privacy notice



(+372) 699-1531

hi@serokell.io

Pille tn 7/5-13, Kesklinna linnaosa, Tallinn, Harju maakond, 10135, Estonia

Serokell: Rate 5.0 based on 9 Google Business reviews

© 2015–2021 Serokell