How can we determine whether a problem is linear/nonlinear?

Before running any classifier just like that, we need to understand what is a linear classification problem. Look at the 2D data below. In data A, blue and green points can be easily classified by drawing a simple line. The line shown is one of the many solutions that are possible and you will get maximum accuracy. A problem is that, if the separating line is drawn improperly, then even in a simple linear case we can mess up. Therefore, proper optimization of the parameters to find an optimum line/plane/hyper-plane is very important. In data B we cannot draw a line that can classify both green and blue points. Therefore, data A represents a linear classification problem and data B represents a non-linear classification problem.

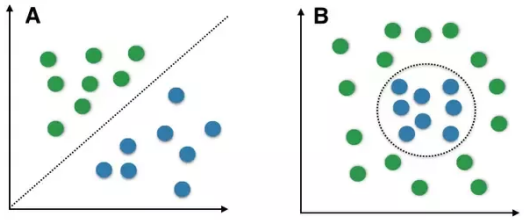

Figure 1

Let's get into details of linear and nonlinear classifiers.

Linear Classifier

Let's say we have data from two classes (o and χ) distributed as shown in the figure below. To discriminate the two classes, one can draw an arbitrary line, s.t. all the 'o' are on one side of the line and χ's on the other side of the line. These two classes are called linearly-separable.

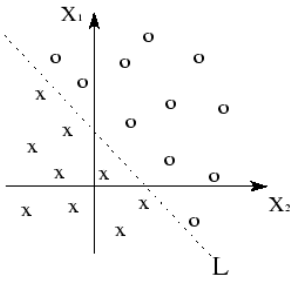

Figure 2

However, there are infinite number of lines that can be drawn to discriminate two separable classes, as shown below

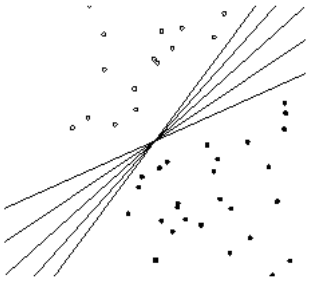

Figure 3

How you approximate the exact location of this discriminating line (or plane or hyperplane) depends on the type of a classifier called linear classifier. Some examples of linear classifier are: Linear Discriminant Classifier, Naive Bayes, Logistic Regression, Perceptron, SVM (with linear kernel).

Non-linear Classifier

Let's consider the famous XOR problem. An XOR's truth table and the data distribution is shown below. In this situation, there is no way that a straight line can be drawn to discriminate the two classes (0 and 1).

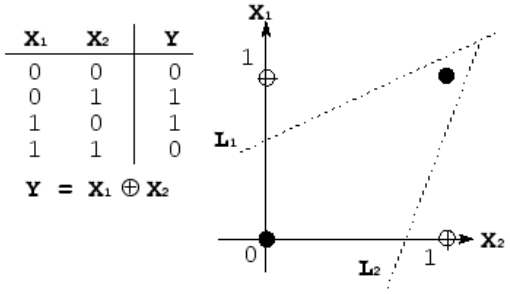

Figure 4

Another example is in the next page, where majority of points one one class can be discriminated from the other classes by drawing a straight line. But some points of that class are present in between the data points of second class (shown inside a circle). It is clear that drawing one straight line cannot identify all the points of one of the class correctly.

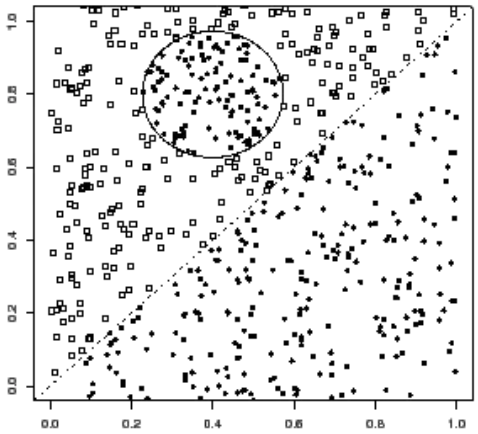

Figure 5

The above problems are called nonlinear classification problems and cannot be solved by drawing a linear classifier; therefore, other alternatives are required. We may need piece-wise linear (i.e. linear in parts), or non-linear classification boundaries to identify the two classes correctly. What would be that non-linear function (or approximation) and how that boundary may look like is defined by a non-linear classifier. Some examples of non-linear classifiers are: Quadratic Discriminant Classifier, Multi-layer Perceptron (most deep networks), Decision Trees, Random Forest, K-Nearest Neighbour.

References:

For more details on Nonlinear Classification can refer below:
`http://www.moseslab.csb.utoronto.ca/alan/Lec6_ML4bio_2012.pdf`