

Support Vector Machine” (SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges. However, it is mostly used in classification problems.

In this we have to find a zoo for every animal given a set of features for the animal.

1. After downloading the dataset using One Hot Encoding we converted categorical variables into a form that could be used by us.
2. SVM kernel is used to take data as an input and transform it into the required form. Different SVM algorithms use different types of kernels.

The kernels used are :

- Linear Kernel
- Radial Basis Kernel
- Polynomial Kernel
- Sigmoid Kernel

1. Linear Kernel

Accuracy: 1.0

It is useful when dealing with large sparse data vectors. It is often used in text categorization. The splines kernel also performs well in regression problems.

$$k(x, y) = 1 + xy + xy \min(x, y) - \frac{x + y}{2} \min(x, y)^2 + \frac{1}{3} \min(x, y)^3$$

2. Polynomial Kernel

Accuracy: 0.9

It is popular in image processing.

$$k(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^d$$

3. Radial Basis Kernel

Accuracy: 0.76

It is general-purpose kernel; used when there is no prior knowledge about the data

$$k(x, y) = \exp\left(-\frac{\|x - y\|}{\sigma}\right)$$

4. Sigmoid Kernel

Accuracy: 0.75

We can use it as the proxy for neural networks.

$$k(x, y) = \tanh(\alpha x^T y + c)$$