DEVYANI BEOHAR

J068

SVR(Support Vector Regression)

Code:

class sklearn.svm.**SVR**(**, kernel='rbf', degree=3, gamma='scale', coef0=0.0, tol=0.001, C=1.0, epsilon=0.1, shrinking=True, cache_size=200, verbose=False, max_iter=- 1*)

kernel*{'linear', 'poly', 'rbf', 'sigmoid', 'precomputed'}, default='rbf'*
Specifies the kernel type to be used in the algorithm.

degree*int, default=3*
Degree of the polynomial kernel function ('poly'). Ignored by all other kernels.

gamma*{'scale', 'auto'} or float, default='scale'*
Kernel coefficient for 'rbf', 'poly' and 'sigmoid'.

- if gamma='scale' (default) is passed then it uses 1 / (n_features * X.var()) as value of gamma,
- if 'auto', uses 1 / n_features.

coef0*float, default=0.0*
Independent term in kernel function. It is only significant in 'poly' and 'sigmoid'.

tol*float, default=1e-3*
Tolerance for stopping criterion.

C*float, default=1.0*

Regularization parameter. The strength of the regularization is inversely proportional to C.

epsilon*float, default=0.1*

Epsilon in the epsilon-SVR model. It specifies the epsilon-tube within which no penalty is associated in the training loss function with points predicted within a distance epsilon from the actual value.

shrinking*bool, default=True*

Whether to use the shrinking heuristic.

cache_size*float, default=200*

Specify the size of the kernel cache (in MB).

verbose*bool, default=False*

Enable verbose output.

max_iter*int, default=-1*

Hard limit on iterations within solver, or -1 for no limit.

The SVM regression algorithm is referred to as **Support Vector Regression** or **SVR**.

Support Vector Regression is a supervised learning algorithm that is used to predict discrete values. Support Vector Regression uses the same principle as the SVMs. The basic idea behind SVR is to find the best fit line. In SVR, the best fit line is the hyperplane that has the maximum number of points.

Unlike other Regression models that try to minimize the error between the real and predicted value, the SVR tries to fit the

best line within a threshold value. The threshold value is the distance between the hyperplane and boundary line. The fit time complexity of SVR is more than quadratic with the number of samples which makes it hard to scale to datasets with more than a couple of 10000 samples.

Advantages of Support Vector Regression

Although Support Vector Regression is used rarely it carries certain advantages that are as mentioned below:

1. It is robust to outliers.
2. Decision model can be easily updated.
3. It has excellent generalization capability, with high prediction accuracy.
4. Its implementation is easy.

Disadvantages of Support Vector Regression

Some of the drawbacks faced by Support Vector Machines while handling regression problems are as mentioned below:

1. They are not suitable for large datasets.
2. In cases where the number of features for each data point exceeds the number of training data samples, the SVM will underperform.
3. The Decision model does not perform very well when the data set has more noise i.e. target classes are overlapping..