

SVM

1. What Are the Basic Assumption?

There are no such assumptions

2. Advantages

1. SVM is more effective in high dimensional spaces.
2. SVM is relatively memory efficient.
3. SVM's are very good when we have no idea on the data.
4. Works well with even unstructured and semi structured data like text, Images and trees.
5. The kernel trick is real strength of SVM. With an appropriate kernel function, we can solve any complex problem.
6. SVM models have generalization in practice, the risk of over-fitting is less in SVM.

3. Disadvantages

1. More Training Time is required for larger dataset
2. It is difficult to choose a good kernel function
3. The SVM hyper parameters are Cost -C and gamma. It is not that easy to fine-tune these hyper-parameters. It is hard to visualize their impact

4. Whether Feature Scaling is required?

Yes

5. Impact of Missing Values?

Although SVMs are an attractive option when constructing a classifier, SVMs do not easily accommodate missing covariate information. Similar to other prediction and classification methods, in-attention to missing data when constructing an SVM can impact the accuracy and utility of the resulting classifier.

6. Impact of outliers?

It is usually sensitive to outliers <https://arxiv.org/abs/1409.0934#:~:text=Despite%20its%20popularity%2C%20SVM%20has,causes%20the%20sensitivity%20to%20outliers.>

Types of Problems it can solve(Supervised)

1. Classification
2. Regression

Overfitting And Underfitting

In SVM, to avoid overfitting, we choose a Soft Margin, instead of a Hard one i.e. we let some data points enter our margin intentionally (but we still penalize it) so that our classifier don't overfit on our training sample

<https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html>

Different Problem statement you can solve using Naive Baye's

1. We can use SVM with every ANN usecases
2. Intrusion Detection
3. Handwriting Recognition

Practical Implementation

1. <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html>
2. <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html>

Performance Metrics

Classification

1. Confusion Matrix
2. Precision, Recall, F1 score

Regression

1. R2, Adjusted R2
2. MSE, RMSE, MAE