

Credit risk assessment Classification model with Support Vector Machines

SELVARAJ X(2577127)

02-NOV-2023

Introduction

- ❖ Credit risk assessment is the process of evaluating the likelihood of a borrower defaulting on a loan. It is a critical function for financial institutions, as it helps them to make informed lending decisions and minimize losses.
- ❖ Support Vector Machines (SVMs) are a type of machine learning algorithm that can be used for both classification and regression tasks.
- ❖ SVMs are well-suited for credit risk assessment because they are able to learn complex relationships between features and outcomes, even with limited data.

Classification with Linear SVM

```
# Linear SVM

from sklearn.svm import SVR
from sklearn.metrics import mean_squared_error, r2_score

support_reg = SVR(kernel='linear')
print('='*100)
print('Hyperparameter:\n', support_reg.get_params())
print('='*100)
# Train phase
support_reg.fit(xtrain, ytrain)
# Testing phase
ypred = support_reg.predict(xtest)
print('Predicted price:\n', ypred)
print('='*100)
```

Evaluate the linear SVM model performance

Accuracy score : 0.134358974358974

classification Report :

		precision	recall	f1-score	support
Confusion Martix :					
[17 21 23 15 17 21 27]	0	0.10	0.12	0.11	141
[27 16 20 16 20 18 20]	1	0.12	0.12	0.12	137
[28 16 19 16 28 25 17]	2	0.13	0.13	0.13	149
[27 18 25 21 16 28 18]	3	0.17	0.14	0.15	153
[14 18 14 22 17 21 15]	4	0.13	0.14	0.13	121
[17 19 19 16 14 24 18]	5	0.16	0.19	0.17	127
[34 24 22 16 21 13 17]]	6	0.13	0.12	0.12	147
	accuracy			0.13	975
	macro avg	0.14	0.14	0.13	975
	weighted avg	0.14	0.13	0.13	975

Classification with Non-linear SVM

- ❖ Non-linear SVMs are needed when the data is not linearly separable. This can happen when the data is noisy or when the relationship between the features and the target variable is complex.
- ❖ The choice of kernel depends on the specific dataset and the problem being solved. Some common kernels include the Radial Basis Function (RBF) kernel and the Polynomial kernel.
- ❖ The RBF kernel is a good general-purpose kernel that can be used for a wide variety of datasets. The Polynomial kernel can be used to model non-linear relationships between the features and the target variable.

Hyperparameter Tuning

- ❖ Hyperparameters are parameters that control the learning process of a machine learning algorithm. They are not learned from the data, but must be set manually by the user.
- ❖ Some important hyperparameters to tune for SVM models include:
 - C: The cost parameter controls the trade-off between maximizing the margin between the classes and minimizing the number of misclassified training examples.
 - Kernel: The kernel function specifies how the data is transformed before being fed to the SVM algorithm. Common kernel functions include linear, polynomial, and radial basis function (RBF).

Decision Boundary Visualization

- ❖ The decision boundary of a classification model is the surface that separates the different classes in the feature space.
- ❖ Linear SVM models produce linear decision boundaries, while non-linear SVM models can produce non-linear decision boundaries.

Support Vectors

Support vectors are the training data points that lie closest to the decision boundary. They are important because they define the decision boundary and have a significant impact on the model's performance.

```
# support vector of Kernal Linear model
```

```
model.support_vectors_
```

```
array([[29. ,  1. , 17. , ..., 39. ,  3. ,  4.1],  
       [42. ,  1. ,  3. , ..., 48. ,  2. ,  3.4],  
       [62. ,  1. , 11. , ..., 37. ,  3. ,  4. ],  
       ...,  
       [56. ,  0. , 15. , ..., 16. ,  1. ,  2.8],  
       [30. ,  1. ,  0. , ..., 41. ,  0. ,  4.7],  
       [37. ,  0. ,  8. , ...,  8. ,  3. ,  3.2]])
```



```
# support vector of Non-Linear Kernal 'RBF' model  
model.support_vectors_
```

```
array([[29. ,  1. , 17. , ..., 39. ,  3. ,  4.1],  
       [42. ,  1. ,  3. , ..., 48. ,  2. ,  3.4],  
       [62. ,  1. , 11. , ..., 37. ,  3. ,  4. ],  
       ...,  
       [56. ,  0. , 15. , ..., 16. ,  1. ,  2.8],  
       [30. ,  1. ,  0. , ..., 41. ,  0. ,  4.7],  
       [37. ,  0. ,  8. , ...,  8. ,  3. ,  3.2]])
```

Model Comparison

- ❖ Linear SVM models are simpler and faster to train than non-linear SVM models.
- ❖ However, non-linear SVM models can be more accurate, especially when the data is non-linearly separable.
- ❖ Linear SVM models are simpler and faster to train than non-linear SVM models. However, non-linear SVM models can be more accurate, especially when the data is non-linearly separable.

Real-World Application

- ❖ Credit risk assessment is used by financial institutions to make a variety of decisions, including:
 - Whether to approve or deny a loan application
 - Setting interest rates
 - Granting overdraft protection
 - Allocating credit limits

Benefits of the Model

Accurate credit risk assessment can benefit financial institutions by:

- ❖ Reducing losses due to defaults
- ❖ Improving portfolio quality
- ❖ Making better lending decisions
- ❖ It can also benefit borrowers by:
 - ❖ Making it easier to obtain loans
 - ❖ Improving creditworthiness
 - ❖ Securing lower interest rates
- ❖ Presentation and Recommendations

Conclusion

- ❖ SVM models are a powerful tool for credit risk assessment.
- ❖ They can be used to learn complex relationships between features and outcomes, and to produce accurate predictions even with limited data.
- ❖ Financial institutions can benefit from using SVM models for credit risk assessment by reducing losses due to defaults, improving portfolio quality, and making better lending decisions.