ort-vector-machines-assignment-3

September 13, 2023

```
# Q1. In order to predict house price based on several characteristics, such as␣
↪location, square footage, number of bedrooms, etc., you are developing an␣
↪SVM regression model.
# Which regression metric in this situation would be the best to employ?
```

In the context of developing an SVM regression model to predict house prices based on various characteristics, you can consider several regression metrics to evaluate the model's performance.

For predicting house prices with an SVM regression model, the best regression metric to employ is typically the Mean Absolute Error (MAE) because it provides a straightforward measure of the average absolute difference between predicted and actual prices, which is easy to interpret in dollars.

```
# Q2. You have built an SVM regression model and are trying to decide between␣
↪using MSE or R-squared as your evaluation metric.

# Which metric would be more appropriate if your goal is to predict the actual␣
↪price of a house as accurately as possible?
```

If your goal is to predict the actual price of a house as accurately as possible, the more appropriate evaluation metric to use is Mean Squared Error (MSE).

MSE measures the average squared difference between the predicted and actual prices. It heavily penalizes larger errors, which means that it provides a strong incentive to minimize prediction errors. In the context of predicting house prices, you want to minimize the error in your predictions to get as close as possible to the actual prices. Therefore, MSE is a suitable metric for this goal.

```
# Q3. You have a dataset with a significant number of outliers and are trying␣
↪to select an appropriate regression metric to use with your SVM model.

# Which metric would be the most appropriate in this scenario?
```

When dealing with a dataset that has a significant number of outliers, it's often more appropriate to use the Mean Absolute Error (MAE) as the regression metric with your SVM model.

Here's why MAE is a good choice in this scenario:

Robustness to Outliers: MAE is less sensitive to outliers compared to Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). Since it calculates the average absolute difference between predicted and actual values, extreme outliers won't disproportionately impact the metric.

```
# Q4. You have built an SVM regression model using a polynomial kernel and are
↪trying to select the best metric to evaluate its performance.

# You have calculated both MSE and RMSE and found that both values are very
↪close.

# Which metric should you choose to use in this case?
```

When you have calculated both Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) and found that both values are very close, it's generally a good practice to choose the simpler and more interpretable metric, which, in this case, is MSE.

MSE is often the simpler and more interpretable choice, even when using a polynomial kernel. It directly measures the average squared difference between your model's predictions and the actual values, making it a straightforward metric to understand and communicate.

```
# Q5. You are comparing the performance of different SVM regression models
↪using different kernels (linear, polynomial, and RBF) and are trying to
↪select the best evaluation metric.

# Which metric would be most appropriate if your goal is to measure how well
↪the model explains the variance in the target variable.
```

If your goal is to measure how well the model explains the variance in the target variable, the most appropriate evaluation metric to use is R-squared (R^2). R-squared, also known as the coefficient of determination, quantifies the proportion of the variance in the target variable that is explained by the model.

if your goal is to measure how well the SVM regression models with different kernels explain the variance in the target variable, R-squared is the most appropriate evaluation metric. It provides a clear and standardized way to assess and compare the goodness of fit across models with various kernels.