**Discussion and Result Comparisons of the two supervised model**

The two supervised models that are used to train the Boston dataset are Linear Regression and Support Vector Machine. The evaluation results are shown below:

|  |  |  |
| --- | --- | --- |
| **Evaluation in Train Dataset** | **Linear Regression** | **Support Vector Machine** |
| RMSE | 5.291267334085863 | 5.363540216530558 |
| R2 score | 0.6493177656003104 | 0.639672474435614 |

|  |  |  |
| --- | --- | --- |
| **Evaluation in Test Dataset** | **Linear Regression** | **Support Vector Machine** |
| RMSE | 6.041044703455954 | 6.1437557911054235 |
| R2 score | 0.6162026947666575 | 0.6030409467174964 |

The provided evaluation results show the performance of two different machine learning models, linear regression and support vector machine (SVM), on both training and testing datasets. Let's discuss and compare these results and then address the pros and cons of each model and conclude with which one to choose overall.

**Comparison:**

RMSE: Both models have relatively similar RMSE values on the testing set, with linear regression having a slightly lower RMSE. This indicates that both models are performing similarly in terms of prediction accuracy.

R2 Score: Linear regression has a slightly higher R2 score on both the training and testing sets, indicating that it explains a bit more of the variance in the data compared to the SVM. However, the difference is not very substantial.

**Pros and Cons:**

*Linear Regression:*

Pros:

Simplicity: Linear regression is simple and interpretable.

Fast Training: Training time is usually fast.

Cons:

Limited Flexibility: It assumes a linear relationship between features and the target variable, which might not hold in all cases.

Sensitive to Outliers: Linear regression can be sensitive to outliers.

*Support Vector Machine:*

Pros:

Flexibility: SVMs can handle both linear and non-linear data relationships through kernel tricks.

Robust to Outliers: SVMs are less affected by outliers due to their margin-based approach.

Cons:

Complexity: SVMs can be computationally expensive and might require more tuning.

Less Intuitive: SVMs are less intuitive and harder to interpret compared to linear regression.

**Overall Choice:**

The choice between linear regression and SVM depends on the specific problem and dataset. In this case, since both models are performing fairly similarly, it might be preferable to choose the simpler model, which is linear regression. It provides similar predictive performance with greater interpretability and less complexity. However, if the data has non-linear relationships that linear regression cannot capture, or if robustness to outliers is crucial, then SVM might be a better choice.

**Analysis and Conclusion:**

In this specific scenario, both models perform similarly in terms of predictive accuracy, with linear regression having a slight edge. However, the choice between the two should consider factors like model interpretability, computational resources, and the nature of the data. It's important to note that model selection should not solely rely on these evaluation metrics; domain knowledge and problem-specific considerations should also play a significant role in making the final decision.