

Checking Linear Regression Fit: Takeaways



by Dataquest Labs, Inc. - All rights reserved © 2024

Syntax

- Predicting values from a fitted model:

```
model.predict(data)
```

- Creating a residuals plot:

```
import matplotlib.pyplot as plt  
plt.(predictions, residuals)
```

- Returning the mean squared error:

```
from sklearn.metrics import mean_squared_error  
mean_squared_error(y_true, predictions)
```

- Returning the root mean squared error:

```
from sklearn.metrics import mean_squared_error  
mean_squared_error(y_true, y_pred, squared = False)
```

- Returning the coefficient of determination:

```
from sklearn.metrics import r2_score  
r2_score(y_true, y_pred)
```

Concepts

- Residual diagnostics are used to check the assumptions that we typically have on the residuals: zero mean and constant variance over the predicted values.
- The mean squared error is also another metric we can use to assess the average amount of squared error in the data. We can take its square root for an idea of the same dimension as the outcome.
- The coefficient of determination is another metric of model fit. It tells us the proportion of variance explained by the model, in terms of the variance of the errors.

References

- [scikit-learn's Linear Regression model](#)
- [scikit-learn's predict\(\) function](#)
- [scikit-learn's mean_squared_error\(\) function](#)
- [scikit-learn's r2_score\(\) function](#)