

**Linear models for regression** : These models assume that the target value is a linear combination of the features, possibly with some noise or error. They can use different loss functions and regularization terms to fit the data and prevent overfitting. Some examples are Ridge, Lasso, ElasticNet, and LinearRegression.

**Linear models for classification** : These models use a linear function to separate the classes, and can also use different loss functions and regularization terms. Some examples are LogisticRegression, SGDClassifier, Perceptron, and PassiveAggressiveClassifier.

**Robust regression** : These models aim to fit a regression model in the presence of corrupt data, such as outliers or errors. They use different strategies to reduce the influence of the outliers on the fit. Some examples are RANSAC, TheilSen, and HuberRegressor.

**Generalized linear models** : These models extend linear models by using a link function to relate the linear predictor to the expected value of the target, and by using a distribution from the exponential family to model the target. They can handle different types of target values, such as counts, probabilities, or positive values. Some examples are PoissonRegressor, GammaRegressor, and TweedieRegressor.

To evaluate the model performance, you can use different metrics that measure how well the predicted values match the actual values. Some of the common metrics for linear regression are:

- **Coefficient of determination or R-squared ( $R^2$ )**: This metric indicates how much of the variation in the target variable is explained by the linear model. It ranges from 0 to 1, with higher values indicating better fit.  $R^2$  can be calculated as  $1 - (\text{sum of squared residuals} / \text{total sum of squares})$ .
- **Root mean squared error (RMSE)**: This metric measures the average magnitude of the prediction errors. It is the square root of the mean of the squared differences between the predicted and actual values. RMSE can be calculated as  $\sqrt{\text{mean}((y_{\text{pred}} - y_{\text{true}})^2)}$ .
- **Mean absolute error (MAE)**: This metric measures the average absolute value of the prediction errors. It is the mean of the absolute differences between the predicted and actual values. MAE can be calculated as  $\text{mean}(\text{abs}(y_{\text{pred}} - y_{\text{true}}))$ .
- **Mean squared error (MSE)**: This metric measures the average squared value of the prediction errors. It is similar to RMSE, but without taking the square root.

MSE can be calculated as  $\text{mean}((y_{\text{pred}} - y_{\text{true}})^2)$ .

You can use the scikit-learn library to implement these metrics in Python. For example, to calculate R2 and RMSE, you can use the following code:

```
# Import libraries
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score, mean_squared_error
import numpy as np

# Create sample data
X = np.array([[0, 0], [1, 1], [2, 2], [3, 3]])
y = np.dot(X, np.array([1, 2])) + 3

# Fit a linear regression model
reg = LinearRegression().fit(X, y)

# Predict on the same data
y_pred = reg.predict(X)

# Calculate R2 score
r2 = r2_score(y, y_pred)
print("R2 score:", r2)

# Calculate RMSE
rmse = np.sqrt(mean_squared_error(y, y_pred))
print("RMSE:", rmse)
```

The output of this code is:

```
R2 score: 1.0
RMSE: 0.0
```

Linear regression is a machine learning algorithm that tries to find a linear relationship between a target variable and one or more independent variables<sup>1</sup>. It can be used for prediction, estimation, or finding the correlation between variables<sup>2</sup>.

To implement linear regression in Python, you can use the scikit-learn library, which provides a `LinearRegression` class that can fit and predict on data<sup>3</sup>. Here is an example of how to use it:

```
# Import the library
from sklearn.linear_model import LinearRegression

# Create some sample data
X = [[1], [2], [3], [4]] # Independent variable
```

```
y = [3, 5, 7, 9] # Target variable

# Create a linear regression object
reg = LinearRegression()

# Fit the model to the data
reg.fit(X, y)

# Predict on new data
y_pred = reg.predict([[5], [6]])

# Print the predicted values
print(y_pred)
```

The output of this code is:

```
[11. 13.]
```

This means that the linear regression model predicts that when X is 5, y is 11, and when X is 6, y is 13. The model has learned the equation  $y = 2X + 1$  from the data.

To evaluate the performance of the linear regression model, you can use different metrics that measure how well the predicted values match the actual values<sup>4</sup>. Some of the common metrics are R-squared, RMSE, MAE, and MSE. You can use the scikit-learn library to calculate these metrics as well. For example, to calculate R-squared and RMSE, you can use the following code:

```
# Import the library
from sklearn.metrics import r2_score, mean_squared_error

# Calculate R-squared
r2 = r2_score(y, y_pred)

# Print R-squared
print("R-squared:", r2)

# Calculate RMSE
rmse = mean_squared_error(y, y_pred, squared=False)

# Print RMSE
print("RMSE:", rmse)
```

The output of this code is:

R-squared: 1.0

RMSE: 0.0

This means that the linear regression model has a perfect fit to the data, as it explains 100% of the variation in the target variable and has no prediction errors.