

However, if the value of λ approaches zero, the regularization parameter has negligible effects on the model, hence resulting in overfitting the model. Regularization is an important technique and should be used when injecting polynomial features into linear or logistic regression classifiers to learn non-linear relationships.

Applying Regularization to Models with Scikit-learn

The technique of adding a penalty to restrain the values of the parameters of the model is also known as Ridge regression or Tikhonov regularization. In this section we will build a linear and logistic regression model with regularization.

Linear Regression with Regularization

This code block is similar to the polynomial linear regression example in Chapter 19. The model will predict house prices from the Boston house-prices dataset. However, this model includes regularization.

```
# import packages
from sklearn.linear_model import Ridge
from sklearn import datasets
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
from math import sqrt
from sklearn.preprocessing import PolynomialFeatures

# load dataset
data = datasets.load_boston()

# separate features and target
X = data.data
y = data.target

# create polynomial features
polynomial_features = PolynomialFeatures(2)
X_higher_order = polynomial_features.fit_transform(X)
```

```

# split in train and test sets
X_train, X_test, y_train, y_test = train_test_split(X_higher_order, y,
shuffle=True)

# create the model. The parameter alpha represents the regularization
magnitude
linear_reg = Ridge(alpha=1.0)

# fit the model on the training set
linear_reg.fit(X_train, y_train)

# make predictions on the test set
predictions = linear_reg.predict(X_test)

# evaluate the model performance using the root mean square error metric
print("Root mean squared error (RMSE): %.2f" % sqrt(mean_squared_error(y_
test, predictions)))

'Output':
Root mean squared error (RMSE): 3.74

```

Take note of the following:

- The method `Ridge(alpha=1.0)` initializes a linear regression model with regularization, where the attribute ‘alpha’ controls the magnitude of the regularization parameter.

Logistic Regression with Regularization

This code block here is also similar to the example in Chapter 20 on logistic regression. The model will predict the three species of flowers from the Iris dataset. The addition to this code segment is the inclusion of a regularization term to the logistic model using the ‘`RidgeClassifier`’ package.

```

# import packages
from sklearn.linear_model import RidgeClassifier
from sklearn import datasets
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score

```