CHAPTER 21

Regularization for Linear Models

Regularization is the technique of adding a parameter, λ, to the loss function of a learning algorithm to improve its ability to generalize to new examples by reducing overfitting. The role of the extra regularization parameter is to shrink or to minimize the measure of the weights (or parameters) of the other features in the model.

Regularization is applied to linear models such as polynomial linear regression and logistic regression which are susceptible to overfitting when high-order polynomial features are added to the set of features.

How Does Regularization Work

During model building, the regularization parameter λ is calibrated to determine how much the magnitude of other features in the model is adjusted when training the model. The higher the value of the regularization, the more the magnitude of the feature weights is reduced.

If the regularization parameter is set too close to zero, it reduces the regularization effect on the feature weights of the model. At zero, the penalty the regularization term imposes is virtually non-existent, and the model is as if the regularization term was never present.

Effects of Regularization on Bias vs. Variance

The higher the value of λ (i.e., the regularization parameter), the more restricted the coefficients (or weights) of the cost function. Hence, if the value of λ is high, the model can result in a learning bias (i.e., it underfits the dataset).

E. Bisong, *Building Machine Learning and Deep Learning Models on Google Cloud Platform*,
https://doi.org/10.1007/978-1-4842-4470-8_21

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)
```