

## CHAPTER 24

# More Supervised Machine Learning Techniques with Scikit-learn

This chapter will cover using Scikit-learn to implement machine learning models using techniques such as

- Feature engineering
- Resampling methods
- Model evaluation methods
- Pipelines for streamlining machine learning workflows
- Techniques for model tuning

## Feature Engineering

Feature engineering is the process of systematically choosing the set of features in the dataset that are useful and relevant to the learning problem. It is often the case that irrelevant features negatively affect the performance of the model. This section will review some techniques implemented in Scikit-learn for selecting relevant features from a dataset. The techniques surveyed include

- Statistical tests to select the best  $k$  features using the **SelectKBest** module
- Recursive feature elimination (RFE) to recursively remove irrelevant features from the dataset