1. Introduction   Print to PDF ▸

The best way to get a sense of the challenges underlying the design of a credit card fraud detection system (FDS) is by designing one. This chapter presents an implementation of a baseline FDS and covers the main steps that need to be considered.

Section 3.2 first describes a simple simulator for payment card transaction data. Although simplistic, the simulator provides a sufficiently challenging setting to approach a typical fraud detection problem. In particular, the simulator will allow to generate datasets that i) are highly unbalanced (low proportion of fraudulent transactions), ii) contain both numerical and categorical variables (with categorical features that have a very high number of possible values), and iii) feature time-dependent fraud scenarios.

Section 3.3 and section 3.4 address the two main steps of a standard predictive modeling process: Feature transformation, and predictive modeling. These sections will provide some baseline strategies for performing meaningful feature transformations, and build a first predictive model whose accuracy will serve as a reference in the rest of the book.

Finally, in section 3.5, we apply this baseline methodology (feature transformation and predictive modeling) to a real-world dataset of card transactions and illustrate its ability to effectively detect fraudulent transactions.