in Chapter 19, but for now it’ll serve as a motivating challenge.) You may be required to
scour vast amounts of data yet have only very few known instances of fraud, so they won’t
be well represented as a class in the training phase. There are also typically fraudulent cases
that escaped detection, which adds to the challenge of training a model to distinguish
between fraud and nonfraud, as those are mislabeled as valid cases. Further, while statisti-
cal outliers in the data—which are signs of unusual activity—are good places to search for
fraud, a typical “alert” system built only on these will have far too many false alarms to be
useful in practice.

Before you begin modeling, successful fraud detection requires looking at the entire busi-
ness process and identifying where fraud can originate. Begin the project with a careful
evaluation of the client’s existing business process. Then collect cases of fraud that have
been found by auditors or others through the existing manual processes. From knowledge
of the business process and these known cases, design metrics for measuring fraud and
work with the client to automate their calculation. Finally, develop the detection models.
This process delivers value to clients at each stage.

The return on investment (ROI) in fraud detection data mining can be extremely impres-
sive. Take the time to develop metrics of success and the baseline (“before”) performance so
that you can score your impact (the “after” picture, which should look better). For example,
on a large project led by one of our colleagues, the client had an alert system for its enor-
mous data processing task whose warnings turned out to be fraud only 1% of the time (very
inefficient, though far better than random). With the data mining solution, however, the hit
rate improved to an astonishing 25%. In another fraud detection project, our colleagues
were able to achieve a savings of over $20 million on an engagement that took less than a
staff-year to complete and deliver. Wouldn’t it have been great if, in the business negotia-
tions, the confident and expert team had been able to negotiate a share of the success? Even
5% \times (after–before) is serious compensation.

Project Methodology and Deliverables

The authors have found that an investment in teamwork and creative problem solving
early in a project pays off down the road. It is essential to cooperate and communicate
closely (both with team members and the client) throughout a project to successfully imple-
ment the solution on time. You will need client experts to convey business understanding
and requirements throughout the project, so get to know them and treat them with respect.
Join client domain expertise with your expertise in business analysis, systems engineering,
and technical skills in predictive modeling.

We recommend operating in a rapid-prototyping framework, where a baseline solution
is developed quickly, to discover the technical and interface issues the final system will
face. Then strengthen the system by iteratively improving components—often possible in
parallel. This allows the path-critical components to be identified and improved as time
and budget allow, and provides decision makers with better estimates of the trade-offs
involved. Include regular administrative and technical updates and on-site meetings at crit-
ical junctures. A client’s great fear is that it will launch analysts on a project which will