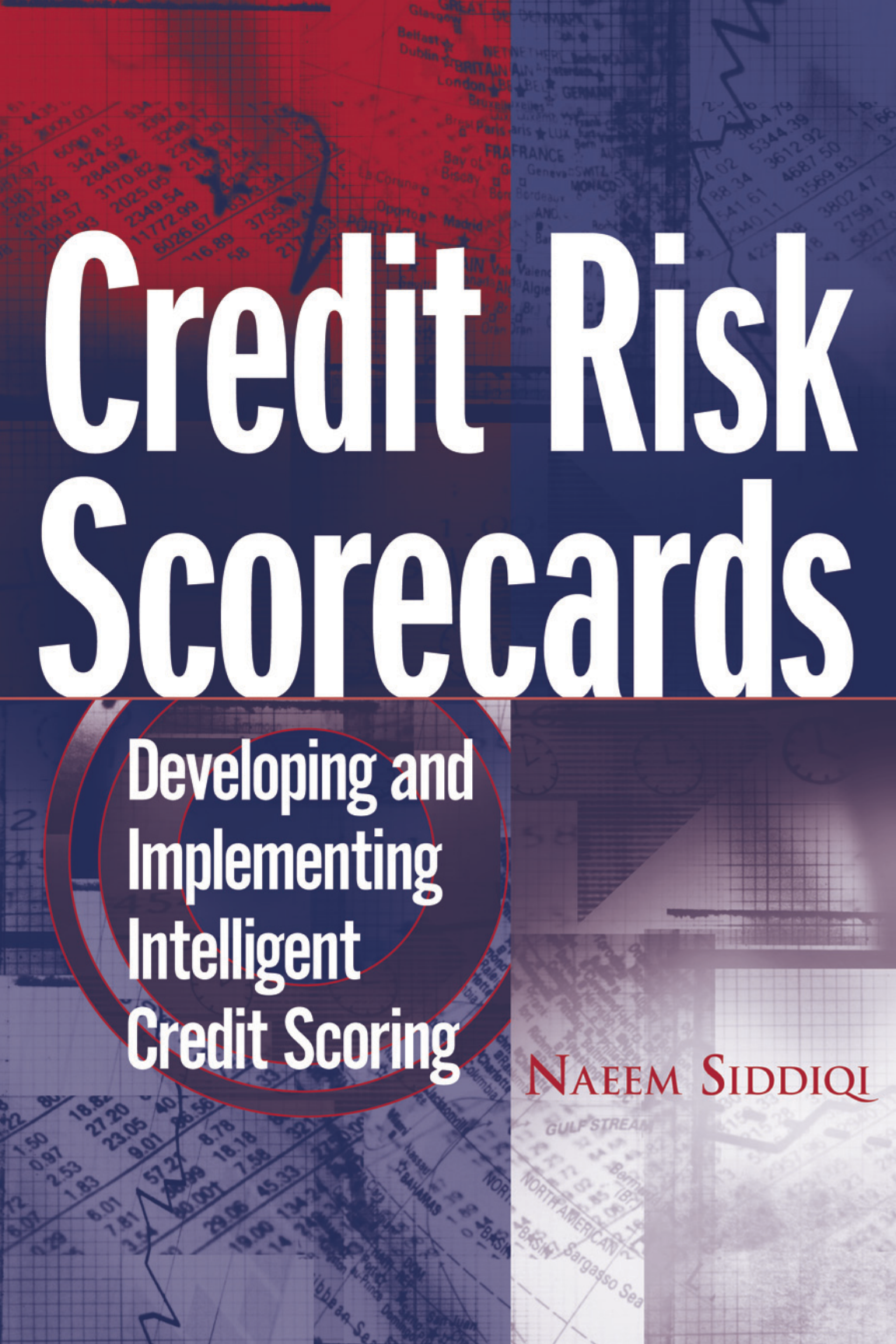
Credit Risk Scorecards
Developing and Implementing Intelligent Credit Scoring
Naeem Siddiqi

Contents

CHAPTER 1

Introduction

Increased competition and growing pressures for revenue generation have led credit-granting and other financial institutions to search for more effective ways to attract new creditworthy customers, and at the same time, control losses. Aggressive marketing efforts have resulted in deeper penetration of the risk pool of potential customers, and the need to process them rapidly and effectively has led to growing automation of the credit and insurance application and adjudication processes. The Risk Manager is now challenged to produce risk adjudication solutions that can not only satisfactorily assess creditworthiness, but also keep the per-unit processing cost low, while reducing turnaround times for customers. In addition, customer service excellence demands that this automated process be able to minimize denial of credit to creditworthy customers, while keeping out as many potentially delinquent ones as possible. In the insurance sector, the ability to keep the prices of policies commensurate with claims risk becomes more critical as underwriting losses increase across the industry.

At the customer management level, companies are striving ever harder to keep their existing clients by offering them additional products and enhanced services. Risk Managers are called on to help in selecting the "right" (i.e., low risk) customers for these favored treatments. Conversely, for customers who exhibit negative behavior (nonpayment, fraud), Risk Managers need to devise strategies to not only identify them, but also deal with them effectively to minimize further loss and recoup any monies owed, as quickly as possible.

It is in this environment that risk scorecards offer a powerful, empirically derived solution to business needs. Risk scorecards have been used by a variety of industries for uses including predicting delinquency nonpayment—that is, bankruptcy—fraud, claims (for insurance), and recovery of amounts owed for accounts in collections. Scoring methodology offers an objective way to assess risk, and also a consistent approach, provided that system overrides are kept to a minimum.

In the past, financial institutions acquired credit risk scorecards from a handful of credit risk vendors. This involved the financial institution providing their data to the vendors, and the vendors then developing a predictive scorecard for delivery. While some advanced companies have had internal modeling and scorecard development functions for a long time, the trend toward developing scorecards in-house has become far more widespread in the last few years. This happened for various reasons.

First, application software became available that allowed users to develop scorecards without investing heavily in advanced programmers and infrastructure. Complex data mining functions became available at the click of a mouse, allowing the user to spend more time applying business and data mining expertise to the problem, rather than debugging complicated programs. The availability of powerful "point and click"–based Extract-Transform-Load (ETL) software enabled efficient extraction and preparation of data for scorecard development and other data mining. Second, advances in intelligent and easy to access data storage have removed much of the burden of gathering the required data and putting it into a form that is amenable to analysis.

Once the tools became available, in-house development became a viable option for many smaller and medium-sized institutions. The industry could now realize the significant Return on Investment (ROI) that in-house scorecard development could deliver for the right players. Experience has shown that in-house credit scorecard development can be done faster, cheaper, and with far more flexibility than before. Development was cheaper, since the cost of maintaining an in-house credit scoring capability was less than the cost of purchased scorecards. Internal development capability also allowed companies to develop far more scorecards (with enhanced segmentation) for the

same expenditure. Scorecards could also be developed faster by internal resources using the right software—which meant that custom scorecards could be implemented faster, leading to lower losses.

In addition, companies realized that their superior knowledge of internal data and business insights led them to develop better-performing scorecards. Defining the population performance definitions is a critical part of scoring system construction, and the ability to vary definitions for different purposes is key. For example, a probability of default score designed for capital planning purposes may exclude moderately delinquent accounts (60 days past due twice during the past 24 months) that are normally included in "bad behavior" and go by the Basel definition for loans considered likely to default (associated with write-off, repossession, foreclosure, judgments, or bankruptcy). This will vary by type of loan or trade line—for example, revolving, installment, mortgage, and so forth. On sample construction, some Scorecard Developers eliminate large numbers of accounts associated with inactivity, indeterminate behavior, and so forth, and this is another area where some empirical investigation and control is warranted.

Better-performing scorecards also came about from having the flexibility to experiment with segmentation, and from following through by developing the optimum number and configuration of scorecards.

Internal scorecard development also increases the knowledge base within organizations. The analyses done reveal hidden treasures of information that allow for better understanding of customers' risk behavior, and lead to better strategy development.

In summary, leaving key modeling and sampling decisions to "external experts" can prove to be a suboptimal route at best, and can also be quite costly. A perfect example that comes to mind is a finance company that outsourced scorecard development and found upon system implementation that the "updated scorecards" turned down 65% of their current and repeat customers, even though they developed specific individual scorecards for present versus former borrowers. Ultimately, the problem was traced back to the good/bad performance definitions and the fact that their average "good" paying customer had delinquency characteristics that would normally be categorized as bad behavior, or indeterminate at the very least! Unfortunately, there were five regional

scorecards for each of the two groups, so that ultimately ten scorecards were shelved at an average cost of $27,000. There was also fallout with customers who were initially turned down after 20 years of doing business with the company.

This book presents a business-focused process for the development and implementation of risk prediction scorecards, one that builds upon a solid foundation of statistics and data mining principles. Statistical and data mining techniques and methodologies have been discussed in detail in various publications, and will not be covered in depth here. The key concepts that will be covered are:

- The application of business intelligence to the scorecard development process, so that the development and implementation of scorecards is seen as an intelligent business solution to a business problem. Good scorecards are not built by passing data solely through a series of programs or algorithms—they are built when the data is passed through the analytical and business-trained mind of the user.
- Collaborative scorecard development, in which end users, subject matter experts, implementers, modelers, and other stakeholders work in a cohesive and coherent manner to get better results.
- The concept of building a risk profile—building scorecards that contain predictive variables representing major information categories. This mimics the thought processes of good risk adjudicators, who analyze information from credit applications, or customer behavior, and create a profile based on the different types of information available. They would not make a decision using four or five pieces of information only—so why should anyone build a scorecard that is narrow-based?
- Anticipating impacts of decisions and preparing for them. Each decision made—whether on the definition of the target variable, segmentation, choice of variables, transformations, choice of cutoffs, or other strategies—starts a chain of events that impacts other areas of the company, as well as future performance. By tapping

into corporate intelligence, and working in collaboration with others, the user will learn to anticipate the impact of each decision and prepare accordingly to minimize disruption and unpleasant surprises.

- View of scorecards as decision support tools. Scorecards should be viewed as a tool to be used for better decision making, and should be created with this view. This means they must be understood and controlled; scorecard development should not result in a complex model that cannot be understood enough to make decisions or perform diagnostics.

Individual scorecard development projects may need to be dealt with differently, depending on each company's unique situation. This methodology should therefore be viewed as a set of guidelines rather than as a set of definitive rules that must be followed. Finally, it is worth noting that regulatory compliance plays an important part in ensuring that scorecards used for granting consumer credit are statistically sound, empirically derived, and capable of separating creditworthy from non-creditworthy applicants at a statistically significant rate.[1] An excellent, but somewhat dated, article on credit scoring and the Equal Credit Opportunity Act is cited in the Bibliography (Hsia, 1978).

Scorecards: General Overview

Risk scoring, as with other predictive models, is a tool used to evaluate the level of risk associated with applicants or customers. While it does not identify "good" (no negative behavior expected) or "bad" (negative behavior expected) applications on an individual basis, it provides statistical odds, or probability, that an applicant with any given score will be "good" or "bad." These probabilities or scores, along with other business considerations such as expected approval rates, profit, churn, and losses, are then used as a basis for decision making.

In its simplest form, a scorecard consists of a group of characteristics, statistically determined to be predictive in separating good and bad accounts. For reference, Exhibit 1.1 shows a part of a scorecard.

EXHIBIT 1.1 SAMPLE SCORECARD (PARTIAL)

Characteristic Name	Attribute	Scorecard Points
AGE	. -> 23	53
AGE	23 -> 25	76
AGE	25 -> 28	79
AGE	28 -> 34	85
AGE	34 -> 46	94
AGE	46 -> 51	103
AGE	51 -> .	105
CARDS	"AMERICAN EXPRESS," "VISA OTHERS," "VISA MYBANK," "NO CREDIT CARDS"	90
CARDS	"CHEQUE CARD," "MASTERCARD/EUROC," "OTHER CREDIT CARD"	99
EC_CARD	0	80
EC_CARD	1	83
INCOME	. -> 500	93
INCOME	500 -> 1,550	81
INCOME	1,550 -> 1850	75
INCOME	1,850 -> 2,550	80
INCOME	2,550 -> .	98
STATUS	"E," "T," "U"	79

Scorecard characteristics may be selected from any of the sources of data available to the lender at the time of the application. Examples of such characteristics are demographics (e.g., age, time at residence, time at job, postal code), existing relationship (e.g., time at bank, number of products, payment performance, previous claims), credit bureau (e.g., inquiries, trades, delinquency, public records), real estate data, and so forth.

Each attribute ("Age" is a *characteristic* and "23–25" is an *attribute*) is assigned points based on statistical analyses, taking into consideration various factors such as the predictive strength of the characteristics, correlation between characteristics, and operational factors. The total score of an applicant is the sum of the scores for each attribute present in the scorecard for that applicant.

Exhibit 1.2 is an example of one of the management reports produced during scorecard development.

The circled line in the exhibit tells us the following:

EXHIBIT 1.2 GAINS CHART

Score Range	Count	Cumulative Count	Number of Goods	Cumulative Number of Goods	Number of Bads	Cumulative Number of Bads	Marginal Badrate	Cumulative Badrate	Approval Rate
273<=Score<279	942	942	940	940	2	2	0.24	0.24	1.91
267<=Score<273	511	1,353	510	1,350	1	3	0.20	0.22	2.91
262<=Score<267	574	1,927	570	1,920	4	7	0.70	0.36	4.14
256<=Score<262	2,087	4,014	2,070	3,990	17	24	0.81	0.60	8.63
250<=Score<256	1,756	5,770	1,740	5,730	16	40	0.91	0.69	12.41
245<=Score<250	2,338	8,108	2,310	8,040	28	68	1.20	0.84	17.44
239<=Score<245	2,917	11,025	2,880	10,920	37	105	1.27	0.95	23.71
233<=Score<239	3,774	14,799	3,720	14,640	54	159	1.43	1.07	31.83
228<=Score<233	2,766	17,565	2,700	17,340	66	225	2.39	1.28	37.77
222<=Score<228	3,366	20,931	3,300	20,640	66	291	1.96	1.39	45.01
216<=Score<222	4,492	25,423	4,380	25,020	112	403	2.49	1.59	54.67
211<=Score<216	4,210	29,633	4,080	29,100	130	533	3.09	1.80	63.73
205<=Score<211	3,455	33,088	3,350	32,460	95	628	2.75	1.90	71.15
199<=Score<205	4,419	37,507	4,260	36,720	159	787	3.60	2.10	80.68
194<=Score<199	1,549	39,056	1,440	38,160	109	896	7.04	2.29	83.99
188<=Score<194	2,006	41,062	1,890	40,050	116	1,012	5.78	2.46	88.31

- For the score range 245–250, the expected marginal bad rate is 1.2%. That is, 1.2% of applicants with a score between 245 and 250 will likely be "bad."
- The cumulative bad rate—that is, the bad rate of all applicants above 245—is 0.84%.
- The acceptance rate at 245 is 17.44%, that is, 17.44% of all applicants score above 245.

Based on factors outlined above, a company can then decide, for example, to decline all applicants who score below 200, or to charge them higher pricing in view of the greater risk they present. "Bad" is generally defined using negative performance indicators such as bankruptcy, fraud, delinquency, write-off/chargeoff, and negative net present value (NPV).

Risk score information, combined with other factors such as expected approval rate and revenue/profit potential at each risk level, can be used

to develop new application strategies that will maximize revenue and minimize bad debt. Some of the strategies for high-risk applicants are:

- Declining credit/services if the risk level is too high
- Assigning a lower starting credit limit on a credit card or line of credit
- Asking the applicant to provide a higher down payment or deposit for mortgages or car loans
- Charging a higher interest rate on a loan
- Charging a higher premium on insurance policies
- Asking the applicant to provide a deposit for utilities services
- Offering prepaid cellular services instead of postpaid
- Denying international calling access from telecommunications companies
- Putting the applicant into a "watch list" for potential fraudulent activity

Conversely, high-scoring applicants may be given preferential rates and higher credit limits, and be offered upgrades to premium products, such as gold or platinum cards, or additional products offered by the company.

Application scores can also help in setting "due diligence" policies. For example, an applicant scoring very high or very low can be declined or approved outright without obtaining further information on real estate, income verification, or valuation of underlying security.

The previous examples specifically dealt with risk scoring at the application stage. Risk scoring is similarly used with existing clients on an ongoing basis. In this context, the client's behavioral data with the company is used to predict the probability of negative behavior. Based on similar business considerations as previously mentioned (e.g., expected risk and profitability levels), different treatments can be tailored to accounts, such as:

- Offering product upgrades and additional products
- Increasing credit limits on credit cards and lines of credit

- Allowing some revolving credit customers to go beyond their credit limits
- Flagging potentially fraudulent transactions
- Offering better pricing on loan/insurance policy renewals
- Deciding whether or not to reissue an expired credit card
- Prequalifying direct marketing lists for cross-selling
- Directing delinquent accounts to more stringent collection methods or outsourcing to a collection agency
- Suspending or revoking phone services or credit facilities
- Put an account into a "watch list" for potential fraudulent activity

In addition to being developed for use with new applicants (application scoring) or existing accounts (behavior scoring), scorecards can also be defined based on the type of data used to develop them. Custom scorecards are those developed using data for customers of one organization exclusively. For example, ABC Bank uses the performance data of its own customers to build a scorecard to predict bankruptcy. It may use internal data or data obtained from a credit bureau for this purpose, but the data is only for its own customers.

Generic or pooled data scorecards are those built using data from multiple lenders. For example, four small banks, none of which has enough data to build its own custom scorecards, decide to pool their data for auto loans. They then build a scorecard with this data and share it, or customize the scorecards based on unique characteristics of their portfolios. Scorecards built using industry bureau data, and marketed by credit bureaus, are a type of generic scorecards.

Risk scoring, in addition to being a tool to evaluate levels of risk, has also been effectively applied in other operational areas, such as:

- Streamlining the decision-making process, that is, higher-risk and borderline applications being given to more experienced staff for more scrutiny, while low-risk applications are assigned to junior staff. This can be done in branches, credit adjudication centers, and collections departments.

- Reducing turnaround time for processing applications through automated decision making
- Evaluating quality of portfolios intended for acquisition
- Setting economic and regulatory capital allocation
- Setting pricing for securitization of receivables portfolios
- Comparing the quality of business from different channels/regions/suppliers

Risk scoring, therefore, provides creditors with an opportunity for consistent and objective decision making, based on empirically derived information. Combined with business knowledge, predictive modeling technologies provide risk managers with added efficiency and control over the risk management process.

In the future, credit scoring is expected to play an enhanced role in large banking organizations, due to the requirements of the new Basel Capital Accord (Basel II). This will also lead to a reevaluation of methodologies and strategy development for scorecards, based on the recommendations of the final accord. In particular, changes may be required in the way "bad" is defined, and in the way the target prediction is connected to "Probability of Default," "Exposure at Default," and "Loss Given Default."

Endnote

1. Regulation B, Section 202.2(p) (1) (i–iv).