Social Media Credit Scoring

Billie Anderson
Ferris State University, USA

J. Michael Hardin
Samford University, USA

INTRODUCTION

Credit scoring is a method of modeling potential risk of credit applicants. It involves using different statistical techniques and past historical data to create a credit score that financial institutions use to assess credit applicants in terms of risk. Credit scoring is essentially a type of classification problem: which credit applicants should be considered good risks and which applicants should be considered bad risks.

A scorecard model is built from a number of characteristic inputs. Each characteristic is comprised of a number of attributes. In the example scorecard shown in Figure 1, age is a characteristic and "25–33" is an attribute. Each attribute is associated with a number of scorecard points. These scorecard points are statistically assigned to differentiate risk, based on the predictive power of the variables, correlation between the variables, and business considerations.

For example, in Figure 1, the credit application of a 32 year old person, who owns his own home and makes $30,000, would be accepted for credit by this institution. The total score of an applicant is the sum of the scores for each attribute present in the scorecard. Lower scores imply a higher risk of default, and higher scores imply lower risk.

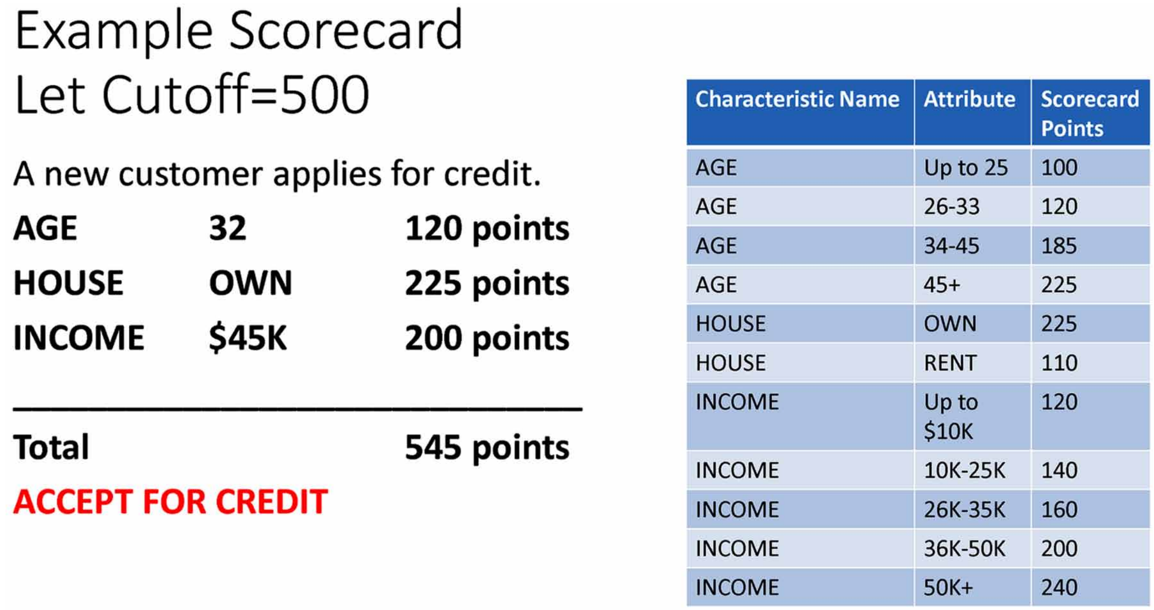
Credit, as it has evolved since the 1950's, it is cold and impersonal, completely based on numbers--what you owe, what you've paid, how much money you have. The numbers all get hacked down to one number: a credit score. The bigger and more complex the global financial system gets, the less it cares about anything other than the applicant's credit score. Banking institutions are becoming less and less personal with online and mobile banking applications. Very few people know the local bank manager at their hometown bank anymore. A credit applicant's personal story is not taken into account when applying for a financial loan. The credit applicant's "character" fits into the bank's algorithms about as well as peanut butter fits into the workings of a Swiss watch.

One consequence of this system is that it is biased against poor people, people with no bank accounts or very little credit history ("thin files") and young people. These types of applicants cannot obtain credit because they cannot generate the financial metrics that would providence evidence to determine if these applicants should be granted or rejected for a financial loan. It is these several billion people who would be the most enthusiastic first customers of a new kind of credit based less on numbers and more on character. The need to market to and provide credit to these types of applicants marks the emergence of a social media credit score.

This chapter will describe how banks and financial organizations are starting to incorporate big data sources, such as data from social media websites, into the credit lending process. A discussion of how more established organizations, such as FICO and SAS, are incorporating big data in their scorecard methodology will be given. A description of two start-up companies

DOI: 10.4018/978-1-5225-2255-3.ch620

Figure 1. Example scorecard



which are using electronic and big data sources such as social media to provide banking services and grant loans will be discussed.

BACKGROUND

The statistical methods used to categorize objects into groups can be traced to 1936 in Fisher's publication (Fisher, 1936). Durand (1941) was the first to use Fisher's methodology to distinguish between good and bad loans. Using this research, the founders of credit scoring, Bill Fair and Earl Isaac, built the first credit scoring system for the United States in 1958. Although credit scoring has been in use since that time, it is only recently that credit scoring has become widespread.

There are some older, yet still very relevant, credit scoring resources that discuss the statistical issues in developing a credit scorecard (Siddiqi, 2006; Thomas, Oliver, & Hand, 2005; Hand & Henley, 1997). These references provide a traditional framework that credit analysts have used in the past to develop credit scores.

As with any other business sector that is highly data dependent, the credit scoring world has felt the impact of the big data phenomenon that is sweeping through modern businesses (Lohr, 2015). Banks are moving away from using traditional statistical techniques to build a credit scorecard. Financial organizations are combining the traditional techniques with new big data technology and big data analytics to build their scorecards. With the rise of companies using social media data as a tool to build better predictive models, financial institutions are integrating big data sources in calculating credit risk (Fei et al., 2015). Recently, even Facebook is hinting at entering the social media credit arena. Facebook has recently made its Messenger application the ability to make financial payments, similar to Apple Pay. What is interesting about Facebook's decision is that it will allow Facebook to collect data on their customer's financial payments and coupled with the already-abundant social media data it has, this decision allows Facebook to become a stand-alone financial lending institution that can incorporate character into the credit score (Maney, 2016).

Wei et al. (2015) provide the most contemporary and comprehensive review of incorporating social media data into an applicant's credit score.