Chapter 4
Predictive Analytics

Sema A. Kalaian
Eastern Michigan University, USA

Rafa M. Kasim
Indiana Tech University, USA

ABSTRACT

Predictive analytics and modeling are analytical tools for knowledge discovery through examining and capturing the complex relationships and patterns among the variables in the existing data in efforts to predict the future organizational performances. Their uses become more common place due largely to collecting massive amount of data, which is referred to as "big data," and the increased need to transform large amounts of data into intelligent information (knowledge) such as trends, patterns, and relationships. The intelligent information can then be used to make smart and informed data-based decisions and predictions using various methods of predictive analytics. The main purpose of this chapter is to present a conceptual and practical overview of some of the basic and advanced analytical tools of predictive analytics. The chapter provides a detailed coverage of some of the predictive analytics tools such as Simple and Multiple-Regression, Polynomial Regression, Logistic Regression, Discriminant Analysis, and Multilevel Modeling.

INTRODUCTION

Predictive analytics, which is also referred to as predictive modeling techniques, are used in a variety of disciplines and fields of study such as business, management, engineering, marketing, technology, actuarial science, information systems, health informatics, and education. Predictive modeling methods are quantitative statistical techniques that are used most often to make future business predictions and decisions based on past historical data (Evans & Lindner, 2012). Kuhns and Johnson (2013) defines predictive modeling as "the process of developing a mathematical tool or model that generates an accurate prediction" (p. 2). Their uses become more common place due largely to

1. Collecting massive amount of data, which is referred to as "big data"; and
2. Increasingly complex nature of related predictive research and problems.

The main objective of predictive modeling is to predict an unknown value of a dependent (outcome) variable from known values of a set of exploratory independent (predictor) variables by analyzing and capturing the relationships between the dependent and independent variables in any research problem (Maisel & Cokins, 2014; Siegel 2014). The results and findings of these analyses are then used to make future predictions such as predicting specific future trends, risks, and behavior patterns. Consumer purchasing patterns, consumer loyalty, credit risks, credit limits, tax fraud, unemployment rates, and consumer attrition are examples of such predictions of future trends and consumer behavior patterns that the businesses and organizations often deal with.

Generally, predictive modeling is a complex data analytic process, of which predictive model building is only a part of the analytic process. The predictive analytics process includes:

1. Understanding the predictive research problem and the massive data to be analyzed.
2. Managing the data, preparing the data for analysis.
3. Analyzing the data and building the analytic models.
4. Evaluating the results and accuracy of the predictive modeling.
5. Deploying and tailoring the final models to directly addressing the original predictive research problem.

One of the most difficult tasks for predictive analysts, researchers and students conducting predictive modeling is identifying the most appropriate predictive analytic technique that can be utilized to answer a particular research oriented business question. In addition, the type and number of the dependent variables are two important primary factors to determine the appropriateness of the predictive modeling technique that can be used for a particular business problem. Provost and Fawcett (2013) stated that the success in today's data-driven businesses requires being able to think about how to correctly apply the principles, concepts, and techniques of predictive modeling to particular predictive business problems.

With the rise of using the internet and electronic devices (e.g., smart phones) to collect massive amount of data as well as technological advances in computer processing power and data storage capabilities, the demand for effective and sophisticated knowledge discovery and predictive modeling techniques has grown exponentially over the last decade. These knowledge discovery and predictive analytics techniques help business executives and policy makers make informed decisions to solve complex organizational and business problems. For example, the survival of businesses and organizations in a knowledge-and-data driven economy is derived from the ability to transform large quantities of data and information to knowledge (Maisel & Cokins, 2014; Siegel 2014). This knowledge can be used to make smart and informed data-based decisions and predictions using predictive analytics techniques. In fact, a decade ago, most such data was not collected or entirely overlooked as a key resource for business and organizational success because lack of knowledge and understanding of the value of such information (Hair, 2007).

Although Business Data Analytics (BDA) is commonly viewed from three major perspectives: descriptive, predictive, and prescriptive (Evans & Lindner, 2012), the focus of this chapter is only on predictive analytics. Therefore, the main purpose of this chapter is to present a conceptual and practical overview of some of the modeling tools and methods that can be used to perform predictive analytics. The conceptual and methodological overview provides analysts with the skills necessary to understand

Related Content

An Entropy-based Mathematical Formulation for Straight Assembly Line Balancing Problem
Ahmad Heydari, Ali Mahmoodiradand Sadegh Niroomand (2016). *International Journal of Strategic Decision Sciences (pp. 57-68).*
www.irma-international.org/article/an-entropy-based-mathematical-formulation-for-straight-assembly-line-balancing-problem/163961

Model of a Performance Measurement System for Maintenance Management
José Contreras, Carlos Parra, Adolfo Crespo Márquez, Vicente González-Prida, Fredy A. Kristjanpollerand Pablo Viveros (2017). *Optimum Decision Making in Asset Management (pp. 194-214).*
www.irma-international.org/chapter/model-of-a-performance-measurement-system-for-maintenance-management/164052

Competitive Advantage: Analysis and Strategy
 (2012). *Systems Thinking and Process Dynamics for Marketing Systems: Technologies and Applications for Decision Management (pp. 251-273).*
www.irma-international.org/chapter/competitive-advantage-analysis-strategy/65309

The Backbone of Decision Support Systems: The Sensor to Decision Chain
Philipp Hertweck, Jürgen Moßgraber, Efstratios Kontopoulos, Panagiotis Mitzias, Tobias Hellmund, Anastasios Karakostas, Désirée Hilbring, Hylke van der Schaaf, Stefanos Vrochidis, Jan-Wilhem Blumeand Ioannis Kompatsiaris (2021). *Research Anthology on Decision Support Systems and Decision Management in Healthcare, Business, and Engineering (pp. 25-47).*
www.irma-international.org/chapter/the-backbone-of-decision-support-systems/282578

Access and Use of Information by Primary Health Care Providers in Rural Uganda: A Qualitative Approach
Maria G. N. Musoke (2010). *International Journal of Decision Support System Technology (pp. 1-9).*
www.irma-international.org/article/access-use-information-primary-health/43906