Predictive Analytics: A Gold-Mine Yet To Be Exploited To Its Zenith

Akanksha Pandey
Information Technology Department,
VESIT,
Mumbai-74,
India.

Abstract

The proliferation, ubiquity and increasing power of computer technology has increased the volume of data collection and it`s storage manifold. This led to continual growth in the size of data sets with consequent increase in complexity as well. Hands-on data analysis is being increasingly augmented with indirect, automated data processing techniquesclustered together and known as DATA MINING.

Primarily, data mining deals with the analysis of data sets for identification of hidden patterns, trends and data values. Data miming in any line of business includes certain common classes of tasks: Anomaly detection (deviation detection), Association rule learning (dependency modelling), Clustering (discovering groups and structures), Classification (applying known structure to new data), Regression (finds a function to model the data with least error), Summarization (compact representation of data set), Forecasting (discovering patterns to make reasonable prediction about the future).The domain of data mining concerned with forecasting is called PREDICTIVE ANALYTICS.

1. Introduction

Today`s mobile technologies and social media have unleashed an exponential increase in information. Predictive analytics, a business intelligence technology is one of the latest to take the future by storm with its immense potential for data- mining and efficacy. Predictive analytics can be defined as any solution that supports the identification of meaningful patterns and correlations among complex, structured and unstructured, historical and potential future data sets for the purpose of predicting future events and accessing the attractiveness of the various courses of action. It is the analytical ability to see relationships between business drivers and performance. Predictive analytics is a relatively new data mining technique which is used to predict future probabilities and trends. The core element in this technique is a 'predictor'. A predictor is a variable around which predictive analytics revolves. A predictor variable can be measured for an individual or multiple entities to predict future behaviour.Predictive analytics encompasses a variety of techniques from predictive modelling, machine learning, and data mining that analyse current and historical facts to make predictions about future, or otherwise unknown, events.

Predictive analytics helps your organization predict with confidence what will happen next so that you can make smarter decisions and improve business

outcomes. Predictive analytics is business intelligence technology that produces a predictive score for each customer or other organizational element. Assigning these predictive scores is the job of the predictive model.

Thus, Predictive Analytics is a more sophisticated form of analytics that is 'forward-thinking' in nature and is used for gaining insights from mathematical or financial modelling. Predictive analytics is being embraced at an increasing rate by organizations that need to gain actionable and forward-looking insight from their data. Predictive analytics simplifies data to amplify value. Predictive analytics increases the precision, accuracy and the speed which an organisation can take decisions on subjects as customers and prospects.

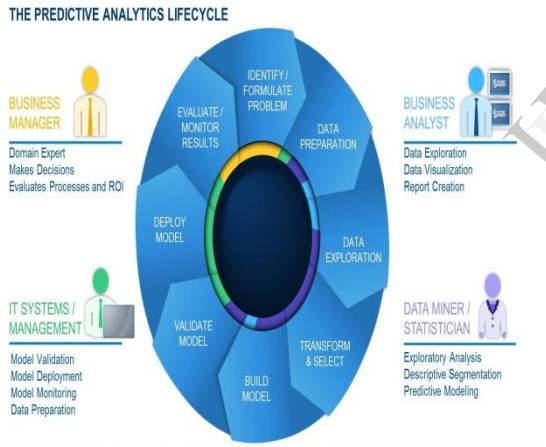

Figure 1. Lifecycle of predictive analytics

2. Predictive modelling

Predictive modelling is a process used to create a statistical model of future behaviour. Multiple predictors are combined into a predictive model, which, when subjected to analysis, can be used to forecast future probabilities with an acceptable level ofreliability. In predictive modelling, an incremental approach is adopted where data is collected, a statistical model is formulated, predictions are made and the model is validated as additional data becomes available. The revised model is subjected to regression testing.The model may employ a simple linear equation or a complex neural network, mapped out by sophisticated software. Predictive models usually operate and perform calculations during live transactions in order to guide decisions. The advancement in computer technology has enabled individual modelling systems to simulate human reaction or behaviour in response to a stimulus or a scenario.

3. Connecting the dots

One of the most serious limitations of traditional predictive analytics was that it could not adapt to changes in environment or subject under consideration. This led to the acceptance of **machine learning** as a part of predictive analytics and thus connecting two domains of technology- analytics and artificial intelligence. Arthur Samuel defined machine learning as a "Field of study that gives computers the ability to learn without being explicitly programmed". Machine learning is a branch of artificial intelligence that enables computers to learn. It helps the system designer to model a predictive system in a more generalized manner. Machine learning deals with processing of large volumes of data, identifying data patterns and instances using algorithms and designing a model in a generalized way which would easily adapt to unseen or unknown data instances. Thus, the core of machine learning deals with representation and generalization. Machine learning techniques emulate human cognition and make use of artificial neural networks to serve the purpose.

Artificial neural networks (ANNs) brought about a revolution in the field of artificial intelligence by making use of multiple chips to simulate a human brain. The problem of the processing capability of a

computer processor not matching that of a human being is not new. As a solution to this problem, models inspired by biological central nervous system came into light. These models were capable of machine learning and pattern recognition, a feature that legacy models lacked or were deprived of. This all new smart data processing model came to be known as artificial neural network. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurones: devices with many inputs and one output) working in unison to solve specific problems. Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by other computer techniques.

4. Tools for predictive analytics

In an attempt to provide a standard platform for predictive analytics Predictive Modeling Markup Language (PMML) was proposed and PMML 4.0 was released in June, 2009. This XML based markup language provides a way to define predictive models using different tools which can also be shared amongst PMML compliant platforms.
Traditionally, predictive analytics tools required expert skills. Moreover, understanding the results delivered by these tools required in –depth knowledge of the same. However, modern predictive analytics tools are no longer restricted to IT specialists. These tools are sophisticated enough to solve, analyse and present meaningful and useful results. Predictive analytics tools range from those that need very little user sophistication to those designed for expert practitioners. The notable open source tools available include Weka, KNIME, Orange, etc. and commercial tools include MATLAB, Angoss KnowledgeSTUDIO, KXEN Modeler, Oracle DataMining, Pervasive, TIBCO, etc.

Figure 2 .Software and tools used for predictive analytics.

5. Applications

On a macro level, organisations need Predictive Analysis for strategic planning, financial planning, focussing on priorities, competitive analysis, achieving profit and revenue targets, developing competitive advantages and differentiation. The uses for predictive analytics are extensive and growing. Some examples include customer churn analysis, predicting insurance fraud and finding patterns in health related data.

5.1. Financial services

A large credit card issuer saw a $6 million profit boost for every million active account by using predictive analytics to assign an optimal credit line for each customer.

5.2. Insurance

A large Brazilian insurer grew net profits by 130% by using predictive analytics in its underwriting to reduce risk and grow revenue from profitable customers.

5.3. Telecommunications

A major global carrier saved $70 million and decreased net bad debtby **25%** in its first year of using analytics based collection solution. Collectors will pin point which accounts will repay the most.

5.4. Healthcare

Predictive analytics is being efficiently used in the healthcare industry to predict the patient volume for healthcare.

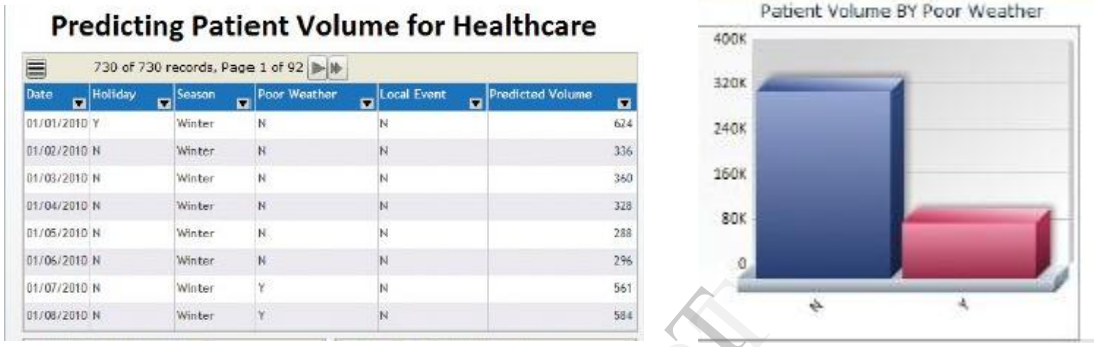

Figure 3. Predicting patient volume for healthcare.

A major collection payer saw more than $20 million in annual savings using an analytics based fraud solution to detect provider fraud and abuse, overpayment, and policy and system errors.

Similar ground-breaking transformation in the airline, transportation and other industries show that predictive analytics can not only improve a single company`s bottom line, it can make an entire industry more efficient.

6. Conclusion

Predictive analytics has immense capability to predict future values of variables with a good deal of preciseness. Therefore, it is used in fields such as meteorology, security, marketing, genetics, economics and many more. Predictive analytics is continuously being used for analysis of various domains and still has great scope for development owing to its efficacy and accuracy.

7. References

[1]Eric Siegel, Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die**,** John Wiley & Sons, 8 March 2013

[2]FICO, Understanding Predictive Analytics: your guide to the science that makes customer decisions smarter, free pdf available: http://www.fico.com/en/Communities/PredictiveAnalytics/Pages/default.aspx

[3]Yet Another Big Data Article, Gerhard Pilcher, Vice President and Senior Scientist, Elder Research, Inc.

[4]Charles Nyce,PhD,CPCU,API,Senior Director of Knowledge Resources, American Institute for CPCU,Insurance Institute of America, Predictive Analytics White Paper, available at: http://www.theinstitutes.org/doc/predictivemodelingwhitepaper.pdf