



The Growing Influence of Predictive Analytics in Various Industries

Prashasthvardhana Prahlad¹, Sachin K. Rai¹, Dr. D.V. Ashoka²

UG Student, Department of Information Science and Engineering, JSSATE, Bengaluru, India¹

Professor, Department of Information Science and Engineering, JSSATE, Bengaluru, India²

Abstract: Predictive analytics is the process of finding structure in historical data to come up with predictions about future events by using a combination of techniques from data mining, machine learning, and statistics. In this process, one usually constructs models to identify relationships among factors that potentially have an impact on future events. Predictive analytics is widely used in marketing, financial services, retail, travel and many other areas. The use of predictive analytics is growing rapidly, which can be attributed to the availability of new data sets and sources, increased computing power, and the constant, unrelenting search for business supremacy. One important fact is that, even though the transactions are taking place in various market sectors, it can be inferred that all the companies has the same strategic goal: to utilize the power of analytics to gain deeper insight into customer demeanor and to use that knowledge to better serve their customers. This paper will give more insight into the rapid expansion of predictive analytics in various markets.

Keywords: Predictive Analytics, Data mining, machine learning, predictive modeling, statistics, markets.

I. INTRODUCTION

Predictive Analytics is a branch of “Advanced Data Analytics” that incorporates many concepts from data mining, machine learning, statistics and predictive modelling, to make certain predictions based on specific patterns and trends observed during the examination and analysis of historical data [2]. Predictive Analytics involves the analysis of transactional data to identify relationships between data sets, and based on other relevant factors, make predictions about the future based on the observed trends [5]. The predictions made in such a manner take utmost importance as they have been derived by the thorough analysis of previous data, and not based on a hunch. By doing so, an organization may identify certain the possible risks and develop a certain counter measure to tackle that risk.

compatible with the selected machine learning method; (3) prepare various new learning models of the modified data to be analysed; (4) derive predictions and report them to the user using the already existing learning model [1].

II. FORMAL DEFINITION

Eric Siegel, a former Columbia University professor and founder of Predictive Analytics World, defines the data analysis method as the power to predict who will click, buy, lie or die [3].

“Predictive analytics is the technology that learns from data to make predictions about what each individual will do — from thriving and donating to stealing and crashing your car. For business, it decreases risk, lowers cost, improves customer service, and decreases unwanted postal mail and spam” - Eric Siegel at an interview regarding the subject [3][4].

III. PREDICTIVE MODELLING

The working of Predictive Analytics can be basically viewed as the combination of advance data analytics and data modelling. Data analytics is used to predict the reliable future trends by developing a predictive model that can predict values for past data or the new data sets that keeps evolving. Data modelling is used for developing a conceptual level data dictionary that can be used effectively by various entities for their respective organisations. Predictive Analytics mainly uses three

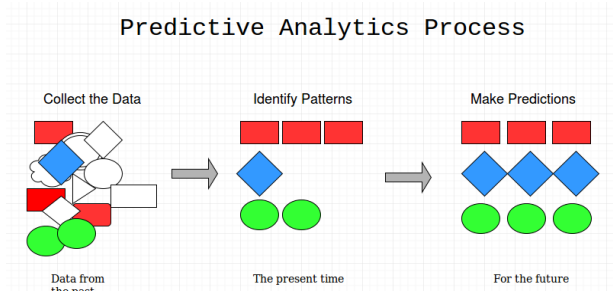


Figure 1: Predictive Analytics Process

Prediction process can be divided into four steps: (1) accumulation and pre-processing of raw data; (2) modification of the pre-processed data into a form that is



modelling techniques, namely, Decision Trees, Regression, and Neural Networks [6]. (1) Decision trees are basically used to classify and partition data in various subsets depending on the different types of input variables. Decision trees provide choice between the numbers of alternatives and also help in identifying the variable that splits the data into different logical groups. Decision Trees are easy to understand and interpret and that's why they are used for preliminary variable selection in modelling. (2) Regression Analysis is used to evaluate relationships among the variables used in data sets. It is used to find key factors in the large data sets that are usually responsible for some crucial event. The main use of Regression Analysis is to predict a response with respect to some event. The response that is depicted is usually restricted as it can only take a limited number of values. (3) the most important and sophisticated of all the modelling techniques is the Neural Networks. These are very popular as they are extremely useful in modelling complex relationships among varied sets of data. They make use of pattern recognition and some artificial intelligence processes to confirm the findings of decision trees and regression analysis. They are very flexible as they don't require mathematical formulas to chart out the relationship between inputs to outputs, making prediction easy.

Other popular techniques employed by Predictive Analytics are: (1) Bayesian Analysis: This analysis is used to define the probability of the degree of belief. (2) Principal Component Analysis: It is used to create principal components i.e. small number of linear combinations that are independent. (3) Time Series Data Mining: It is a combination of traditional data mining and forecasting techniques, which uses techniques such as sampling, clustering and decision trees to the data that has been collected over time and that keeps evolving. This helps in improving the reliability and accuracy of the predictions [6].

IV. INFLUENCE OF PREDICTIVE ANALYTICS IN VARIOUS INDUSTRIES

The main advantage of predictive analytics is that any industry that is looking to minimize the risks, optimize their operation and increase their revenue, can integrate predictive analytic methods into their business model [14]. Currently, the use of predictive analytics is most visible in the following industries: (1) Banking and Financial Services; (2) Healthcare; (3) Tourism and Travel; (4) Manufacturing.

A. Banking and Financial Services

One of the main reasons why Predictive Analytics becomes really important is because of its way of predicting reliably and accurately. Let's take a look at the case of financial sectors, for example, banks. According to the Annual Verizon Data Breach Report [9], most of the data breach cases that take place affects the financial institutions like banks. Recently in India, a major data breach occurred which compromised almost 3.2 million debit cards, leaking account information and also unauthorised use of financial assets of the respective accounts. According to SISA Report, the actual breach occurred on 21st May, 2016 and security agencies came to know about it somewhere in October 2016, spanning almost five months of delay which aggravated the whole situation [8]. SBI one of the major lenders in the country had to start one of the biggest card replacement drives in the country, replacing almost 6lacs debit cards. Predictive analysis can play a crucial role in solving this problem by analysing the cause of similar data breaches that have occurred in the past, like in the Indian case where malwares were used to compromise the Payment Gateway of Hitachi Payment Systems [7]. Using Data Aggregation and Data Co-relation Techniques the past data can be used efficiently to predict the novel security instalments that need to be brought up in the existing security arrangements. This can be supplemented with automated analysis of correlated events and production of alerts which is a reasonably fast way to notify recipients of the immediate issues. Applications can be employed to automate the gathering of data that can expedite the whole process. Predictive analytics can be efficiently used to

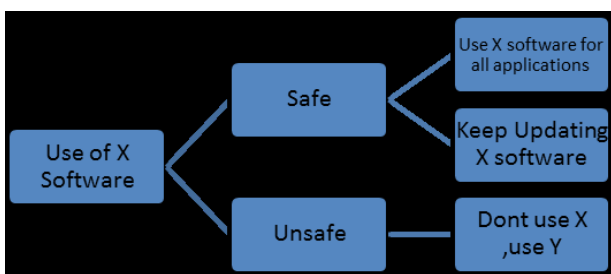


Figure 2: Decision Trees

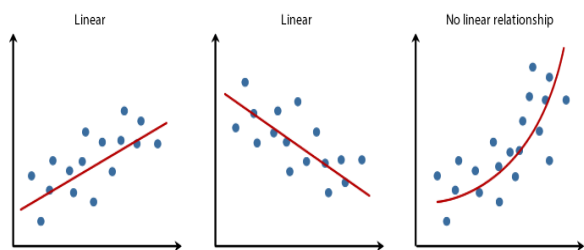


Figure 3: Regression Analysis

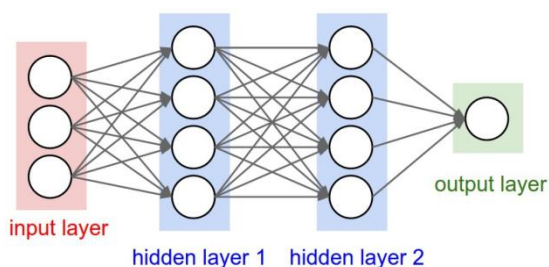


Figure 4: Neural Network



achieve customer relationship management by using analytics to gain an insight about customer's life cycle, their activities trend etc. It can further help in expediting Fraud Detection and Risk Management by predicting accurately and reliably any possible future attacks on the security system. Since the main culprits of various data breaches can be attributed to phishing, hacking, malware injections, there is a need to find novel and better security solutions to tackle the ever changing new ways of breaching the data. This may include the organization to strengthen their network firewall or by developing applications that detect any foreign code. By doing so, the organization is reducing the time between the data breach detection and the actual occurrence. This may save potentially up to billions of customer's money in banks.

B. Healthcare Industry

Predictive Analytics is finding one of its major applications in the healthcare industry. Although primarily it is used for determining what sort of disease a patient might contract in the future, the application of predictive analytics isn't limited to that. The predictions one can make in the healthcare industries include the reactions of different patients to various medications [10]. Other examples include predicting infections, making an assumption as to whether a particular disease might be contracted, helping a physician with the diagnosis and treatment prescription, and even predicting the future wellness of a patient [10].

By using data mining and using the data in several machine learning concepts, healthcare industries have broadened their use of predictive analytics to various other aspects of the healthcare industry such as improving the patient care, chronic disease management, hospital management, and supply chain efficiencies [11].

Broadly, one can summarize the advantages of using predictive analytics in healthcare as follows: (1) Predictive Analytics helps the physicians to give a better diagnosis to the patient by employing several predictive algorithms. For example, for a patient who has been experiencing common flu-like symptoms like cough and cold, the doctor might not know right away if it is just a common-cold or something life threatening. As the predictive algorithms are generated by studying similar cases in the past, it assists the physician in rendering a more accurate diagnosis of the patient. (2) Predictive Analytics help in the administration of preventive medicines. Based on the models that have been generated by predictive algorithms, one can recognize what disease an individual might contract at a certain age and accordingly the medicines can be provided to the patient. This greatly helps reduce the cost of the treatment and also increases the survival chance of a patient. One such example is the various vaccinations that are given to babies as they grow. The antibodies that fight against some major diseases such as chickenpox are administered to children before they reach 13 years of age. This helps the patient to prepare if he/she is ever

diagnosed with the same disease. (3) Predictive analytics provides employers and hospitals with predictions concerning insurance and other healthcare costs. This can be done by deriving an algorithm based on the work pattern of its workers, to make predictions about the future medical costs. This can be explained in another manner, predictive analytics help health insurance providers to provide a plan that is more suitable for a patient with particular characteristics. For example, an insurance plan for a footballer would sound more attractive to the customer if there were any special plans for him/her particularly regarding leg-injuries. It can also be said that predictive analytics can be used to treat different patients in the most appropriate manner, based on the models generated by studying similar medical cases. (4) Patients have an increased likelihood of better end-results due to predictive analytics. The use of predictive analytics will no doubt lead to improvements in the quality of life of patients. Individuals will receive treatments that will work for them, and not be given unnecessary medications just because that medication works for the bulk of the population. The patients also have the chance to be better informed regarding their health due to various wearable devices that utilize predictive algorithms such as the fitness watch, heart-rate monitor and so. All this help the patients to be better-informed regarding their health, and with consultations with their doctors, can make more qualified decisions regarding their lifestyles [10]. The conclusion that can be drawn is that predictive analytics will continue to be extensively used in the healthcare industries. For example, a doctor using predictive analytics can predict that his patient might get a heart attack after 3 years. On hearing this, the patient can make necessary changes in his/her lifestyle so as to reduce the risk of him/her ever getting one in the future.

C. Tourism and Travel Industries

In the travel industry, predictive analytics has a large amount of untapped potential. The huge amount of data, combined with predictive modelling, unlocks a whole new domain of possibilities for airlines, airports, travel agencies and the travellers [12].

Some of the examples are as follows: (1) Recommendations for travel products. A simple predictive algorithm to predict which hotel a particular customer might choose can be predicted based on the previous choice of hotels of particular customer, considering its star rating, price, the brand and other relevant parameters. A more complex algorithm may take into account the duration of his previous stays, and suggest a certain category of hotels for a longer stay and a different type for shorter stays. As for the hotel agencies, they can use predictive analytics to identify various times of the year where the tourist intake is very high and accordingly vary the prices. Also, based on previous tourist turnout, hotels can recognize the seasons when the turnout is less, and take measures such as providing special discounts for



customers who wish to travel during these times. Further, using special algorithms based on their browsing history and other interests, specific offers can be made available only to certain group of customers. For example, for a major football fan, viewing the 2018 World Cup final live in Moscow might be the ultimate dream. Hence, the hotels and airline companies can target such customers and provide them with reduced accommodation and flight rates during that season. (2) Segmentation and clustering of passengers. A basic segmentation of travellers is based on their trip purpose (leisure or business). This is a reliable parameter to explain customer behaviour (price sensitive or quality sensitive) and then used to adapt travel products accordingly. Supervised classification algorithms can be used to segment people in well-known classes. (3) Fraud detection. The number of air passengers exceeds 3 billion per year worldwide. This generates a large number of online transactions that need to be validated in real-time. The losses coming from fraud in payments and other related transactions are significant. Predictive Analytics plays an important role in detecting frauds, and can also be used to detect cyber-attacks on cyber systems. The main challenge on these kinds of applications is to balance the trade-off between the number of “true” detected issues against the “false positive” cases. An example for such “false positive” cases are instances of credit card rejections while the genuine cardholder is travelling abroad. Such important exceptions should be taken into account to better the existing algorithms. (4) Passenger and other travel data enrichment. Analysis of records coming from different travel systems have shown that data is not as rich as we would like. Similarly there is other important information that is not coming from raw data and is available only by integrating other data sources. Predictive analytics is useful to infer missing data and also matching different sources bringing new capabilities to travel systems [12][13].

D. Manufacturing Industries

Traditional analytical models in manufacturing industries have been around for a very long time. One could also find several guess-based forecast models in the same industry. In spite of predictive analytics gradually making its way to manufacture-only companies, there aren't any high profile stories [15]. This can be attributed to the fact that the derivations made using predictive analytic methods are company specific. The methods are such that it mostly works only for that company's processes and products [15][16]. Despite this, we can broadly categorize the advantages of predictive analytics in the manufacturing industry into 4 groups [14]. (1) To improve the quality of the product. This is the fundamental objective of predictive analytics. Databases are aggregated conveniently, redundant data is removed frequently and the data storage is more efficient than ever. Moreover, typical predictive analytics software is propelling towards a less technical analysis by almost entirely automating

these processes. As a result, the overall quality of the predictive analytics model is enhanced, furnishing a more robust plan-of-action for the manufacturer. (2) Demand forecast. Once a product is manufactured, it must be sold. It is found that more often than not, the demand is cyclical or seasonal. In such cases, knowing the peaks, valleys, and short-term changes in demand are the main criteria for resource allocation in manufacturing. Although considering these factors is a staple in any successful manufacturing company, predictive analytics allows manufacturers to consider previously hidden factors. For example, to predict future sales based upon the past sales, predictive analytics takes historical sales data and applies forms of regression. Further, to generate an even better model, predictive analytics can be used to find the factors contributing to those sales, and integrate these parameters into a future model. (3) Machine Utilization. Manufacturing engineers spend much of their time increasing the value of equipment in the factory. Predictive analytics helps with this by automating much of the previously performed analysis. Also, to utilize the machines at the factory in an efficient manner, there is a need for proper machine scheduling. The application of predictive analytics in machine scheduling combines the prognosis for product demand with the products that are currently being manufactured to optimize machine utilization. Usage of predictive analytics also increases the accuracy. (4) Preventive Maintenance in the Factory.

Predictive analytics help to determine the uptime of a machine. By doing so, it saves the company any damage-repair costs that occur due to the untimely and sudden breakdown of any machine in the factory. By studying the history of breakdown of the machines present at the factory, and also combining the patterns found with the signals sent by the sensor in the machines, predictive analytics can detect certain events that can lead to the breakdown of the machine. It is important to note that this application is a new area to manufacturers and is gradually gaining a lot of traction [15]. One must note that the aforementioned application of predictive analytics in manufacturing industries is only a broad grouping. In reality, each company has its own specific needs and varying parameters are to be considered. Therefore, for a manufacturing company to successfully have predictive analytics integrated into its business model, it is essential to consider the type of product it manufactures, the machine and equipment it has, and the type of workforce it employs.

V. CONCLUSION

Predictive analytics is being used in all companies that seek to optimize their operation costs and maximize the revenue can be generated. It can be said that irrespective the type of industry, predictive analytics is the chief method to carry out the business in a profitable manner.



However, care must be taken to ensure that suitable methods and concepts of predictive analytics are utilized at the appropriate industries. This paper has given an insight as to how predictive analytics is currently being used in the banking industry, the healthcare industry, the tourism industry and the manufacturing industry. Despite already being used in a large scale, the progress of predictive analytics is showing no signs of slowing down and will continue to grow as everyone are looking for ways to optimize their current business as well as to maximize the profits in the future.

- [16] Vin Vashishta “Predictive Analytics in Manufacturing: Why and How for Non-Data Scientists” on April 19, 2015. Available from: <https://www.linkedin.com/pulse/predictive-analytics-manufacturing-why-how-non-data-vin-vashishta>

REFERENCES

- [1] Nischol Mishra, Dr. Sanjay Silakari “Predictive Analytics: A Survey, Trends, Applications, Opportunities & Challenges” in International Journal of Computer Science and Information Technologies (IJCSIT), Vol 3 (3), 2012
- [2] Rao T. V. N, Shaik S. A, Kaur S. M. Predictive Analytics - The Cognitive Analysis. Orient.J. Comp. Sci. and Technol; 10(1). Available from:<http://www.computerscijournal.org/?p=5006>
- [3] Chad Brooks “What is Data Analytics?” in Business News Daily, August 4, 2013. Available from: <http://www.businessnewsdaily.com/4938-predictive-analytics.html>
- [4] Ned Smith “Predictive Analytics Unlocks Big Data” in BusinessNewsDaily, April 12, 2013. Available from: <http://www.businessnewsdaily.com/4326-predictive-analytics-unlocks-big-data.html>
- [5] <http://www.predictiveanalyticstoday.com/what-is-predictive-analytics/>
- [6] SAS “Predictive Analytics: What it is and why it matters” Available from: https://www.sas.com/en_us/insights/analytics/predictive-analytics.html
- [7] Final Investigation Report of Hitachi Data Breach. Available from: https://www.hitachi-payments.com/src/HPY%20Press%20Release_V9.pdf
- [8] Audit Report: Malware caused India’s biggest debit card breach. Available from: <http://www.livemint.com/Industry/jVF2Aw72w0DcBsUGseV0UP/Malware-caused-Indias-biggest-debit-card-fraud-Audit-repor.html>
- [9] Annual Verizon Report “Top 9 Causes of Data Breaches”. Available from: <http://www.crn.com/slide-show/security/300076528/verizon-report-top-9-causes-of-data-breaches.htm/pgno/0/7>
- [10] Linda A. Winters-Miner, PhD “Seven Ways Predictive Analytics Can Improve Healthcare” on 6/10/2014. Available from: <https://www.elsevier.com/connect/seven-ways-predictive-analytics-can-improve-healthcare>
- [11] Health Catalyst “Defining Predictive Analytics in Healthcare”. Available from <https://www.healthcatalyst.com/predictive-analytics>
- [12] Rodrigo Acuna Agost “5 examples of Predictive Analytics in the travel Industry” in the Amadeus blog. Available from: <http://www.amadeus.com/blog/07/04/5-examples-predictive-analytics-travel-industry/>
- [13] Rodrigo Acuna Agost “5 more examples of Predictive Analytics in the travel Industry” in the Amadeus blog. Available from:<http://www.amadeus.com/blog/14/04/5-examples-predictive-analytics-travel-industry-2/>
- [14] Jacob LaRiviere, Preston McAfee, Justin Rao, Vijay K. Narayanan, Walter Sun “Where Predictive Analytics is Having its Biggest Impact” in Harvard Business Review on May 25, 2016. Available from: <https://hbr.org/2016/05/where-predictive-analytics-is-having-the-biggest-impact>
- [15] David Gillman “4 Ways Predictive Analytics is Transforming Manufacturing”, April 14, 2014. Available from: <http://it.toolbox.com/blogs/inside-erp/4-ways-predictive-analytics-is-transforming-manufacturing-60819>