

SURVEY ON PROBABILITY ANALYSIS OVER TOOLS AND TECHNIQUES

Suman Antony Lasrado¹, Vishal Dhawadkar² & Kevin Milan Sequeira³

Abstract - Since predictive analysis is advanced part of data engineering it is used to predict some situation based on data. Data mining methods are used by predictive analysis to make prediction about future events. The process include analysis of historic data. A model can be created to predict using Predictive Analysis modelling techniques. Predictive Various statistical & analytical techniques are used to develop models that predict future occurrence ,events or probability for analysis. Predictive analysis is able to not only deal with continuous changes, but discontinuous changes as well. Predictive analysis can generate valuable information for the management of a supply chain company to improve decision-making.

Keywords –Predictive Analysis, Data Mining

1. INTRODUCTION

Predictive analysis is group of two words predict & analysis, but it works in reverse first analyse then predict. It is human nature that we want to know and predict what the future event is. It deals with the prediction of future events based on previously observed historical data. Prediction is the data analysis it can be used to predict future data trends. The prime element of predictive analysis is the „predictor“. It is a variable that can be measured for an individual or entity to predict its future behavior. For example, a credit card company consider age, income, and credit History as predictors to determine the risk factor in issuing a credit card to a customer. In this paper, we examine disaster management systems from a software architecture perspective. We identify the types of systems and the phases within which they are expected to operate. We outline the vital quality attributes required in the different types of software systems developed for disaster management. We examine some of the existing systems developed for responding to emergency and disaster situations. We describe the software architectural concerns addressed by the existing systems. We thus develop a framework (prototype) within which emergency and disaster can be analyzed.

Emergence of big data systems has also made the predictive analysis to grow more prominence. As enterprises have started using larger and broader pools of data in Hadoop clusters and other big data platform opportunities for mining the data for predictive insights have also increased. predictive analysis capabilities have also increased by IT vendors as they have heightened development and commercialization of machine learning tools

2. PROCESS OF PREDICTIVE ANALYTIC

Predictive analytics requires a high level of expertise with statistical methods and the ability to build predictive data models.

2.1 Define Project

Through predictive analysis it is possible to define the deliverables, scope of the effort, outcomes, business objectives, identify the data sets that are going to be used

2.2 Data Collection

Through predictive analysis it is possible to Define the project outcomes, scope of the effort, business objectives, deliverable, identify the data sets that are going to be used.

2.3 Data Analysis

Data Analysis is the process of inspecting, cleaning and modelling data with the objective of discovering useful information, arriving at conclusion

2.4 Statistics

Statistical Analysis enables to validate the assumptions, hypothesis and test them using standard statistical models.

¹ Asst. Professor, St Aloysius College, AİMIT Mangalore, Karnataka, India

² MCA V Semester, St Aloysius College, AİMIT, Mangalore, Karnataka, India

³ MCA V Semester, St Aloysius College, AİMIT, Mangalore, Karnataka, India

2.5 Modeling

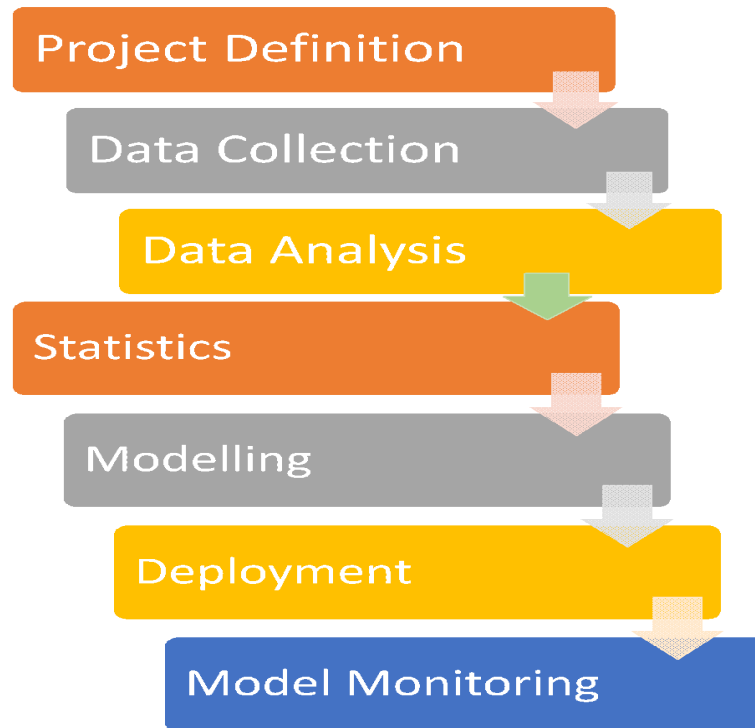
Predictive modelling provides the ability to automatically create accurate predictive models about future. There are also options to choose the best solution with multi-modal evaluation.

2.6 Deployment

Predictive model deployment provides the option to deploy the analytical results into everyday decision making process to get results, reports and output by automating the decisions based on the modelling.

2.7 Model Monitoring

Models are managed and monitored to review the model performance to ensure that it is providing the results expected.



3. PREDICTIVE ANALYSIS TECHNIQUES

The available methods to examine predictive analytics can be categorized into regression and machine learning techniques. It has become easy to store, process huge amounts of data both structured and unstructured with Central Processing Unit, low-priced memory and new technologies like Hadoop, MapReduce, Text analytics and Memory database. This helps to discover unknown patterns and it provide new insights.

3.1 Regression Technique

Regression models are the mainstay of predictive analytics. The focus lies on establishing a mathematical equation as a model to represent the interactions between the different variables in consideration. Depending on the situation, there are a wide variety of models that can be applied while performing predictive analytics. Linear regression model finds the association between a dependent variable y and more than one independent variable X . Logic regression method converts data about the binary dependent variables into a boundless continuous variables. Discrete choice models explains, and guess choices between two or more discrete alternatives. Probit regression model permits the dependent variable can take two values for example male or female

3.2 Machine Learning

Machine learning, a branch of artificial intelligence, was originally employed to develop techniques to enable computers to learn. Today, since it includes a number of advanced statistical methods for regression and classification, it finds application in a wide variety of fields including medical diagnostics, credit card fraud detection, face and speech recognition and analysis of the stock market. In certain applications it is sufficient to directly predict the dependent variable without focusing on the underlying relationships between variables. In other cases, the underlying relationships can be very complex and the mathematical form of the dependencies unknown. For such cases, machine learning techniques emulate human cognition and learn from training examples to predict future events.

4. TOOLS FOR PREDICTIVE ANALYSIS

In the previous section, we categorized disaster management systems. Certain architectural concerns are vital for the development of these systems. We now define from, the meaning of each of the quality attributes and their implication in the context of disaster management.

4.1 R Language

R is a platform for statistical calculations and graphics that runs on a wide variety of Windows, UNIX, and Mac OS Platforms. R provides an extensive range of statistical functionalities such as linear, non-linear modeling, statistical tests, classification and clustering models. It is extremely extensible and provides capabilities for data handling, calculation, and graphical display, calculations on array, and tools for data analysis, programming language that includes loops, conditionals and many other features

4.2 Weka

Weka is a java written group of machine learning algorithms for data mining process. This algorithm may be applied directly or indirectly to a dataset with the help of java code. Weka comprises tools for clustering, association, regression and data preprocessing. In this environment for developing a new machine learning algorithms

4.3 OpenNN

OpenNN is a software library written in the C++ programming language which implements neural networks, a main area of deep learning research. The software implements any number of layers of non-linear processing units for supervised learning. This deep architecture allows the design of neural networks with universal approximation properties. Additionally, it allows multiprocessing programming by means of OpenMP, in order to increase computer performance.

4.4 KNIME

the Konstanz Information Miner, is an open source data analytics, reporting and integration platform. KNIME integrates various components for machine learning and data mining through its modular data pipelining concept. A graphical user interface allows assembly of nodes for data preprocessing (ETL: Extraction, Transformation, Loading), for modeling and data analysis and visualization without, or with only minimal, programming. To some extent KNIME can be considered as an SAS alternative

5. APPLICATION OF PREDICTIVE ANALYSIS

In this section we are going to discuss about the various application that use predictive analysis on day to day basis and make changes accordingly,

5.1 Banking & Financial Services:

The financial industry, with huge amounts of data and money at stake, has long embraced predictive analytics to detect and reduce fraud, measure credit risk, maximize cross-sell/up-sell opportunities and retain valuable customers. Commonwealth Bank uses analytics to predict the likelihood of fraud activity for any given transaction before it is authorized – within 40 milliseconds of the transaction initiation

5.2 Fraud Detection:

Fraud is a big problem for many businesses and can be of various types: inaccurate credit applications, fraudulent transactions (both offline and online), identity thefts and false insurance claims. These problems plague firms of all sizes in many industries. Some examples of likely victims are credit card issuers, insurance companies, retail merchants, manufacturers, business-to-business suppliers and even services providers. A predictive model can help weed out the "bads" and reduce a business's exposure to fraud

5.3 Customer retention:

One of the biggest risks to companies and their growth is the silent customer. These customers do not reach out to companies and let them know if they are dissatisfied with their products or services. With predictive analytics, subtle signals such as specific behaviors and decreasing engagement can easily be spotted well before the customer churns. Once these customers are identified early, the company can then take a full suite of steps to influence the customer's decision to leave.

5.4 Risk Management:

When employing risk management techniques, the results are always to predict and benefit from a future scenario. The capital asset pricing model (CAP-M) "predicts" the best portfolio to maximize return. Probabilistic risk assessment (PRA) when combined with mini-Delphi techniques and statistical approaches yields accurate forecasts. These are examples of approaches that can extend from project to market, and from near to long term. Underwriting (see below) and other business approaches identify risk management as a predictive method.

5.5 Hospital Management:

Experts use predictive analysis in health care primarily to determine which patients are at risk of developing certain conditions, like diabetes, asthma, heart disease, and other lifetime illnesses. Additionally, sophisticated clinical decision support systems incorporate predictive analytics to support medical decision making at the point of care this help the hospitals in properly understanding the patients situation and to provide a necessary medication

6. CONCLUSION

This paper tells about the necessary details of the predictive analysis .This paper will also discuss about the process that a predictive analysis will go through so that it can predict about the results as close as it can, in this paper we will also discuss about the various techniques that the predictive analysis uses in order to get the result and also go through some of the tools that are being used in order to create a predictive analysis .The above mentioned tools are all open source software and not commercial predictive analytic tools and finally we will see about some of the sectors where the predictive analysis are being used

7. REFERENCES

- [1] Predictive analytics , https://www.sas.com/en_us/insights/analytics/predictive-analytics.html
- [2] Predictive analytic for customer retention <https://the-modeling-agency.com/predictive-analytics-for-customer-retention/>
- [3] Predictive Analysis https://en.wikipedia.org/wiki/Predictive_analytics#Technology_and_big_data_influences
- [4] Krishna Gandhi, "A Survey on Predictive Analysis", http://ijirt.org/master/publishedpaper/IJIRT144126_PAPER.pdf
- [5] Predictive Analysis Tools , <https://www.predictiveanalyticstoday.com/predictive-analytics-tools/>