Journal of Contemporary Issues in Business and Government Vol. 28, No. 04, 2022
https://cibgp.com/

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2022.28.04.071

AN OVERVIEW OF TRENDS AND TECHNIQUES IN PREDICTIVE ANALYTICS

Naga Sujana Kakumanu
Research Scholar, Department of GSBH, GITAM (Deemed to be University), Hyderabad Campus
nkakuman@gitam.in
Dr. P. Sridhar
Assistant Professor, Department of GSBH, GITAM (Deemed to be University), Hyderabad Campus
spinnint3@gitam.edu
Bhargavi Mullapudi
Research Scholar, Department of GSBH, GITAM (Deemed to be University), Hyderabad Campus
bmullapu@gitam.in

ABSTRACT

The term "predictive analytics" refers to statistical and analytical methods. This phrase was created using statistics, machine learning, database methods, and optimization techniques. Its roots may be traced all the way back to classical statistics. It creates forecasts using both current and historical data. Predictive analytics algorithms may be used to predict unpredictable behavior and future occurrences. To assign a score, predictive analytics approaches will be employed. A larger number suggests a greater chance of an event occurring, whereas a lower number indicates a lower chance. These frameworks tackle a range of commercial and scientific problems by analyzing historical and transactional data trends. These models assist in identifying the risks and opportunities faced by each customer, employee, or management in a business. As interest in decision-support solutions has grown, predictive analytics models have risen to the top. In this article, we'll go through the methodology, techniques, and applications of predictive analytics.
Keywords: Predictive Analysis, Methodologies, Applications

1. Introduction:

Various tools to be used for knowing predictions. It gives solution for the future prospects. Businesses may successfully use big data to their advantage by using predictive analytics. Businesses may benefit from it by being more proactive, forward-thinking, and predicting trends or data behavior analysis. Its development has coincided with the advancement of big data technology.

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2022.28.04.071

Predictive analytics has a wide variety of uses in e-commerce. Insurance firms acquire data on employees from a trustee, determine which sorts of employees may be engaged in certain private insurance, and then approach them to persuade them to buy their products. Predictive analytics models are used by banks to identify credit card risks and fraudulent clients so that they can be notified. Financial investing firms identify stocks that have the potential to provide a strategic advantage, and they can even estimate future stock value based on previous and current versions. When investing in manufacturing, many other companies utilize predictive algorithms to estimate product sales.

Pharmaceutical companies may identify medications with low sales in a particular area and be notified when such drugs expire.

Process of Predictive Analytics:

Following shows the process.

2.1 Gathering Requirements:

Before developing a model, it is important to define the importance of forecast. The sort of information that will be gathered will be defined by the prediction. Customers meet with data professionals to know about the prerequisites for building a predictive model and how these forecasts would benefit them. For the model to be developed, the customer data will be required.

2.2 Gathering Of information:

After learning about the design specifications, the analyst will collect data from a variety of sources to build the model. Data requires information of consumers regarding usage of the company's goods. The data might be organized or unorganized. Then it is required to checks information from the customers.

2.3 Data manipulation and assessment:

By evaluating and organizing the data, data analysts prepare it for evaluation and usage in the model. This procedure transforms unstructured information into an organized format. When all of the data is in a logical order, it can be evaluated for quality. We must fix both of these concerns if the primary dataset includes inaccurate data or many missing values for the attributes. The analysis step, often known as data management or massaging, includes converting raw data into a format that can be analyzed.

2.4 Machine Learning, Statistics:

Predictive analytics may benefit from a variety of machine learning and statistical approaches. Regression and Probability are the most successful techniques in analytics. Several predictive analytics jobs need machine learning technology such as support vector machines, decision trees, and artificial neural networks. All predictive analytics models are based on machine learning approaches and statistical. As a consequence, the analysts construct prediction models using statistical and machine learning concepts.

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2022.28.04.071

2.5 Modeling for Prediction:

Following creation, the model is tested on a subset of the primary dataset to check its validity, and if effective, the model is fit. The model may provide accurate predictions based on fresh data supplied into the system as input after it has been fitted. In many situations, the multi-model technique aids in the resolution of a problem.

2.6 Forecasting and surveillance:

Predictive analytics isn't a special process in creating future predictions. It's a methodical approach. It's a multi-step process that starts with project planning and continues with rollout and tracking to guarantee that the technology is used efficiently in the legal system.

2. Prospects in Predictive Analytics:

Despite the fact that predictive analytics has a long history and has been widely used in a variety of areas for decades, businesses are continually turning to predictive analytics to enhance their results and profits as a result of technology developments and data dependence.

The following are some of the reasons for this appeal:

- Predictive analytics is used to gather inferences from massive datasets based on the amount and kind of data.
- Some machines are quicker, inexpensive, and quicker to use when it comes to processing.
- There is a vast selection of development tools, and advancements in user-friendly software are still being made.

Because of the data into useful information of user-friendly and dynamic technology, predictive analytics is no longer confined to statisticians and mathematicians. Corporate strategic planners utilise it significantly in their decision-making processes. The following are a few of the most common forecasting options:

3.1 Fraud Detection:

The identification and mitigation of unlawful personality traits can be improved by combining many analytic techniques. Concerns have been raised about the rise of data protection. Behavioral analytics may be used to maintain genuine monitoring of communication channels. It could be able to spot suspicious behavior that can also escalate to a hoax. This method can also be used to block malicious.

3.2 Risk Mitigation:

Using predictive analytics, a credit record can estimate the possibility of a purchaser or user of a provider failing in the future, taking into account all key data around a person's employment. This is how payment card companies and insurance companies identify dishonest consumers.

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2022.28.04.071

3.3 Effectiveness of Marketing Campaigns:

Predictive analytics may be used to anticipate the people will respond to purchasing. I might also like to draw attention to intra- and inter opportunities. It helps businesses sustain a very high - value customers.

3.4 Operational Enhancement:

Merchandise prediction and financial allocation can both benefit from the use of predictive models. Predictive analytics may be used by airlines to determine airfares. By predicting the number of guests on a specific occasion, businesses may use predictive models to optimize vacancy and revenue. Predictive analytics may help a business run more efficiently.

3.5 Clinical Decision-Making Aid:

Decision support systems can be used to rule out other conditions based on prediction models. It can also aid in the development of illness drugs.

4. Predictive Analytics Models Classifications:

Predictive modelling is a wide definition of predictive analytics that refers to evaluate and then estimate outcomes. All analytics-related domains, machine learning is one of these areas, and it is used in business choices.

4.1 Predictive Model:

To investigate the link among the functionality of a mechanism and its attributes. In business, where customer assessment answers are necessary, this model allows the idea how an equal team in a sample choice will display a certain attitude. It imitates social actions in order to respond to a specific query. While the customer is engaging, it calculates the hazards connected with the customer or action.

4.2 Descriptive Model:

The interpretive framework specifies the structure of the data to analyze clusters in a potential. It targets one consumer or accomplishment, but explanatory models discover many links between a goods and its users. Rather of assessing consumers based on their actions, it categorizes them based on their user experience. Many distinct people can be produced another forecast in this model.

4.3 Decision Model:

The pattern is described as a judgment model that predicts the result of a multi-variable choice. These models are used to optimize, maximize a certain result, and minimize a different impact. It's used to create business measures to guarantee the correct action is performed for each customer and scenario.

4.4 Outlier's Model:

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2022.28.04.071

The outlier's model known as data entries in a collection that are beyond the usual. It may recognise outstanding numbers on their own or in conjunction with other figures and groups. For predictive analytics, the outlier model is notably useful in retail and banking.

5. Predictive Analytics Techniques:

Regression analysis is one of the most often used statistical approaches for evaluating the connection between variables. It explains the frequency with which something occurs and one or more variable. It looks at how the value of the factors vary.

5.1 Linear / Non-linear Regression Analysis:

In the case of continuous data with a normal distribution, linear regression is used to find the principal things and to know how certain factors influence a variable's movement. A predictor variable in the regression is used to predict the value of a response variable. In this example, all of the independent factors are mapped to the dependent variables using a regression algorithm. The regression function's prediction of the dependent variable using a probability distribution.

5.2 Time series Analysis:

With time as the input parameter, the time series model consists of a sequence of data points. It creates a numerical meter based on data from the past year and uses that measure to anticipate data for the next three to six weeks. This model's application cases include the number of daily respondents followed in the past three months, sales over the previous 20 quarters, and the number of patients that showed up at a certain hospital in the preceding six weeks. It's a strong method for examining the evolution of a particular variable with an amount of accuracy that goes beyond simple averages. It also takes into account the periods of the year as well as any occurrences that may have an influence on the metric.

5.3 Clustering Analysis:

Clustering organizes data into discrete, tiered inventive groups based on similar qualities. Consider the case of an eCommerce shoe company that wishes to execute personalized marketing campaigns for its consumers. In such scenario, they may have to comb through hundreds of thousands of records to come up with a personalized approach for each consumer. They may simply split clients into comparable groups based on shared features and create approaches for each group on a larger scale using the clustering process.

5.4 Neural Network Analysis:

Artificial neural networks are available in a number of different models, each with its own algorithm. The hidden layer is a well-known technique that is frequently used in supervised learning. Artificial neural networks can also help with unsupervised learning. Clustering is an unsupervised learning approach that incorporates artificial neural networks. They were capable of handling non-linear data relationships. They're also used to evaluate regression and decision tree model outputs.

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2022.28.04.071

5.5 Bayesian Statistics:

The word "degree of belief" is used to measure the chances of things occurring in this method, which is part of statistics. The Bayes theorem, which specifies prior and posterior events, is the foundation of Bayesian statistics. Bayesian network, which is a probabilistic graphical model and to determine the causes and effects of a problem.

5.6 Support Vector Machine:

It is a classifier that sorts samples into groups using a hyperplane. It displays cases in a plane where a noticeable gap divides the examples into separate groups. The new models are then classified into one of two groups, depending on which side of this issue they are on.

5.7 Factor Analysis:

Factor Analysis is a grouping of Principal Component Analysis. Both methods try to estimate the covariance matrix, but factor analysis looks for a preset pattern in the data. The goal of factor analysis is to describe the covariance relationships between numerous variables in terms of a few underlying and non-observable random components known as factors.

5.8 A/B Testing:

A/B testing is a basic prospective randomized experiment. It's a technique for comparing two different versions of a variable in a controlled environment to discover which one performs better. A/B Testing and Predictive Analytics complement each other. A combination of both research techniques is the most effective elixir for optimizing.

5.10 Survival Analysis:

Survival analysis is a form of statistical analysis that examines the impact of predictors on the length of time it takes for a process to occur rather than the probability of it occurring.

6. Predictive Analytics in Action:

Predictive analytics has several applications in a variety of disciplines. We've compiled a list of some of the most popular applications below.

6.1 Financial Services and Banking Sector:

In the banking and financial industries, predictive analytics is widely employed. In all organizations, data and money are essential components, and obtaining insights from that data and money flow is a must. Predictive analytics can help identify customers and transactions. When these firms lend money to their customers, it lowers their credit risk. Cross-selling and up-selling opportunities, as well as customer retention and attractiveness, are all aided by it.

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2022.28.04.071

6.2 Insurance and Health Sector:

In the pharmaceutical business, predictive analytics is utilised to enhance drug formulation and supply chain management. These companies may use this method to predict when drugs will expire in a specific region due to a lack of sales. Predictive analytics algorithms are used to identify and forecast customer fraud claims. Business employs this tool to identify customers who are most likely to acquire a illness and contact them with offers of plans that are the best value for their money.

6.3 Retail Sector:

Predictive analytics may help a retailer discover customers and determine what they want and desire. Using this approach, they can predict how customers will react to a product. Predictive analytics might also help the retail business optimize its supply chain.

6.4 Oil and Gas Sector:

In the oil and gas sector, predictive analytics is used to predict equipment failure and decrease risk. These models are used to estimate future resource needs. To avoid a fatal mistake in the future, energy-based businesses can foresee necessity restoration.

6.5 Government and Public Sector:

Organizations employ big data-based tools to identify possible criminal activities in each location. They utilize to take more precautions for unlawful activities not to happen. Governments utilizes predictive analytics to forecast population changes at the national and state levels. In order to increase safety, predictive analytics methods are widely used.

7. Conclusion:

The use of predictive models for prediction tasks has a long history. Models were formerly used as predictive models, and they were based on big data sets. As computer science and computer methods have improved, more approaches have been developed, and better and better algorithms have been proposed. Because of developments in artificial intelligence and machine learning, intelligent computing techniques and algorithms have changed the world of computation.

References:

1. Eric Siegel, 2016, “Predictive Analytics”, John Willey and Sons Ltd
2. M Schiff, 2012, “BI Experts: Why Predictive Analytics Will Continue to Grow”, The Data Warehouse Institute.
3. V Dhar, 2001, “Predictions in Financial Markets: The Case of Small Disjuncts”, ACM Transaction on Intelligent Systems and Technology, Vol-2, Issue-3.
4. J Feblowitz, 2013, “Analytics in Oil and Gas: The Big Deal About Big Data”, Proceeding of SPE Digital Energy Conference, Texas, USA.

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2022.28.04.071

5. G H Kim, S Trimi, J-H Chung, 2014, "Big-data applications in the government sector", Communications of the ACM, Vol-57, Issue-3, Pages-78-85.
6. J S Armstrong, 2012, "Illusions in regression analysis", International Journal of Forecasting, Vol-28, Issue-3, Pages-689-694.
7. Peter M Lee, 2012, "Bayesian Statistics: An Introduction, 4th Edition", John Willey and Sons Ltd
8. Ben Hur et al, 2001, "Support Vector Clustering", Journal of Machine Learning Research, Vol-2, Pages125-137.
9. J Lin, E Keogh, S Lonardi, C Chiu, 2003, "A symbolic representation of time series, with implications for streaming algorithms", Proceedings of the 8th ACM SIGMOD workshop on research issues in data mining and knowledge discovery, Pages-2-11.
10. H Abdi, L J Williams, 2010, "Principal component analysis", WIREs: Computational Statistics, Vol-2, Issue-4, Pages-433-459.
11. K Das, GS Vidyashankar, 2006, "Competitive Advantage in Retail Through Analytics: Developing Insights, Creating Values", Information Management.
12. N Conz, 2008, "Insurers Shift to Customer-Focused Predictive Analytics Technologies", Insurance & Technology.
13. J Feblowitz, 2013, "Analytics in Oil and Gas: The Big Deal About Big Data", Proceeding of SPE Digital Energy Conference, Texas, USA.
14. G H Kim, S Trimi, J-H Chung, 2014, "Big-data applications in the government sector", Communications of the ACM, Vol-57, Issue-3, Pages-78-85.