A VIEW INTO THE WORLD OF PREDICTIVE ANALYTICS

Ms. Jalajakshi V, Research Scholar, CSE Department Navkis College of Engineering, Hassan jalajakshi.prakash@gmail.com
Dr. Myna A N, Research Supervisor, Associate Professor, Department of CSE, Navkis College of Engineering, Hassan mynaan@gmail.com

ABSTRACT :
Predictive analytics is a word that is mostly used in relation to statistical and analytical procedures. Analytics, deep learning, storage databases approaches, and optimization algorithms are all used to create this phrase Predictive modelling is a more complex area of data engineering is the application data to forecast an event or likelihood. Predictive analytics makes forecasts about future instances using data analysis techniques and then recommendations made based on those forecasts. Its origins can be traced back to probability and statistics. It makes predictions based on data and information. Predictive and prescriptive algorithms can be used to forecast the future events and changeable behaviour. Predictive analytics methods are generally used to assign a score. A higher value obtained in the score suggests a greater likelihood of occurrence, even though a lower figure implies a lower risk. These programs solve a number of commercial and research challenges by analysing historical and operational trend of the data. These models are extremely helpful in identifying hazards and possibilities for each particular consumer, employee, or supervisor in a business. As interest in decision-making services and support has expanded, advanced analytical models have triumphed in this field. In this research, we will look at the methodology, methodologies, and applications of predictive modeling.
Keywords: Data Science, Mathematical modeling formatting, Term dependency modeling, Term weighting, Distributed cluster, Indexing.

INTRODUCTİON :

To foresee upcoming scenarios, advanced predictive modelling, a subgroup of analysis tools, is being applied. It gathers data, data mining, deep learning, and machine intelligence to study historical data in order to make future predictions [1]. It combines new technologies, model construction, and administration with forecasting. By efficiently adopting data analytics, businesses may economically leverage massive data. It could help businesses become more involved, look - ahead, and predictive of data-driven trends or behavior. It has grown in leaps and bounds in unison with the advancement of big data technology. Assume ABC Inc. is an example of an e-commerce corporation. The organization performs its selling operation on a global scale over the internet and sells a wide range of items. Millions of customers visit ABC's webpage to look for a product and its benefits. They look for information about the product's characteristics, pricing, and special offers on ABC's website. The season has an impact on the sales of a variety of goods. For example, growth of air conditioner increases during the summer, whereas demand for geysers increases during the winter. Consumers are seeking things that are appropriate for the season. Here, the ABC Company can collect all of a client's search history, such as the things they are interested in at any given time, and a detailed analysis of the customer's purchasing behaviours. The target price in which a specific buyer is engaged. How people are enticed when they notice special discounts on a product. What additional goods do buyers purchase in addition to one product? XYZ Organization for sure would be incorporating advanced analysis techniques to have predict the advanced consumer demand based on the massive information obtained, they use very highly computational logic programmers to draw patterns for the consumer data, this helps the business of the company to get aligned with the consumer marked in this world. It will determine which specific clients will be enticed by whatever form of advice and then contact them via emails and messages, companies store data and try to attract consumers by a method of their search to attract a similar set of items. They will notify the

client because such an offer is available on the items the customer has on its website. If the customer returns to the website to buy the item, the company will advertise the other items that were sold in connection with it to other customers. If a customer purchases a product on a frequent basis, the company may reduce the offer or increase the price for that customer. This is only one example of the various uses of predictive analytics. In the retail sector, prescriptive modeling has a variety of applications. It has a wide range of applications in a variety of fields. Insurance firms gather expert data from a private organisation, determine which types of working professionals could be interested in specific different insurance plans, and then pursue them to persuade them to buy their products. Forecasting are used by banking institutions to detect credit or debit card risks and dishonest customers, and to alert them[3]. Benefit people in stock portfolios can choose firms that offer a high return on investment and even forecast future market performance using historical and current data. If they spend money on manufacturing, other companies employ predictive models to forecast how well their items will sell[4].

PREDICTIVE ANALYTICS PROCESS :

Prescriptive modeling entails numerous procedures that allow a data scientist to forecast the world based on data and information. The predictive analytics method is depicted in figure 1 below.

Fig. 1. Predictive Modelling Process

REQUIREMENT COLLECTION :

While creating a prediction system, must first be determined whatever the goal of forecasting is. The sort of information that will be gathered should be determined through prediction A drug manufacturer, for example, seeks to investigate how likely it is that a drug will be sold in a specific area in order to avoid the medication's expiration. Consumers are met with data professionals to confirm that the predictive model is needed and that the consumer will benefit from the projections. Which client data will be required for proposed model will be determined?

DATA COLLECTION :

After understanding about the objectives of the customer organisation, the researcher will collect the data, which may come from a variety of sources, that will be used to build the prediction model. This could be a comprehensive list of customers who use or inspect the company's goods. This data could be managed or unmanaged. On their own facilities, the researcher double-checks the data gathered from clients.

DATA ANALYSIS :

Data analysts arrange the obtained data for the analysis and usage in the system by analysing it. This phase converts unorganized data into structured data. Since all of the data is structured and available, its quality is assessed. There is a chance that the primarily affect has erroneous data and that there are multiple missing values for the characteristics; all of this must be handled. The accuracy of the forecasting model is fully dependent on data quality. The analytical phase is also known as data munging or data massaging, and it refers to the process of converting raw data into an analysis-ready format.

STATISTICS, MACHINE LEARNING :

Many Data analysis and ML related high computational modelling techniques are used in the data analysis approach. The most significant approaches utilised in analysis are probabilistic theory and linear regressions. Likewise, machine learning methods such as convolutional neural networks, decision branching trees or random forests and support vector machine (svm) are commonly employed in various data analysis jobs. All predictive analytics methods are having a basic foundation of statistical measures which are quite usefull and machine learning in nature. As a result, analysts use statistics and machine learning techniques to create predictive model. These methods are seen in various industries which changes the impact of today's working model.

PREDICTIVE MODELLING :

During this predictive modelling stage, a model is constructed using machine learning (ML) approaches and the sample database is considered for its future analyzing . futuristic process are taken and it really is evaluated on the test dataset, which would be a subset of the main gathered dataset, to ensure the model's correctness, and if effective, the system has been said to be matched. Once the models is trained, the model may generate correct estimates based on fresh data fed into the system as input. A multi-model solution is chosen for a problem in many applications. In a predictive modelling system, there are various models that can be analyzed using classification, regression techniques. this model has various types of outcomes which can calculated using the accuracies of the model in the application.

PREDICTION AND MONITORING :

After the predictive modelling phase is successfully completed and tested, the model is implemented at the customer's location with high precision and accuracy for daily estimations and decision - making process. Neither the predicted system generates the outcome results not the managerial process involved in this complete process. The model was analyzed on a regular basis to ensure that it is giving recommendation model and correct forecasts. We've demonstrated that predictive analytics isn't simply one step towards making future forecasts. It is a technique that entails numerous steps, from gathering requirements to deployment and monitoring for optimal system performance, as well as rendering it a decision-making system.

PREDICTIVE ANALYSIS ADVANTAGES :

We have witnessed the change in the way the technologies have been drastically changing over time, Despite the fact that predictive modeling has a longstanding experience and it has been extensively implemented in various disciplines for decades, today in this technology era of predictive analytics due to technological advances and reliance on data [5]. Many businesses are turning to predictive analytics to boost their end result and to increase their revenue which helps them to gain profits which can't be possible in normal methods. This appeal can be attributed to a number of factors, including:

1. We can see a rapid plummet in the various types of data that has led to the use of predictive analysis which assist to find valuable insights from a huge size of datasets.
2. This method seems to be a better faster, better, cheaper which helps to be more user-friendly computers which are available for processing.
3. In today technology world, we can see the developments in the number of software's that are being developed every day, we can see a number of advanced technologies developing day by day which makes the common man's life easier.
4. In this evolving competitive world where the technology companies are having cutthroat competition to sustain to get the maximum benefit for the users.

With the development of a wide range of user-friendly and participatory technology, predictive analysis is no longer confined to statistics and mathematics. This is being utilised extensively by

industry professionals like business analysts to take strategic business decisions. Many of the most prevalent advancement in the area of predictive analytics are as follows:

- **Identifying illegal activities**: Integrating numerous analytic approaches helps increase detection and prevention of illegal patterns of behavior. The rise of cybercrime is raising alarm on safety concerns in the society. Behavioral analysis may be used to continuously monitor activity in real time. It may detect unusual actions that may lead to fraud. This technique may also be used to detect threats [6].
- **Mitigating Risk**: The credit rating, using data analytics, may estimate the probability of a purchaser or customer of a trying to do a defaulting activity. The prediction algorithm generates the credit rating by analysing all of the data related to the person's trustworthiness. Credit card firms and insurance providers use this to identify fraudulent consumers [7].
- **Marketing Strategies**: Predictive analytics may be used to forecast customer reactions to product purchases, having a huge database of the customer activities for purchasing, our highly computational algorithms will try to promote a similar kind of products to the users which helps to improve the marketing strategies of the companies It assists companies in attracting and retaining the most lucrative consumers [8].
- **Analysis in Daily operation**: Predictions may be used to predict inventories and manage the resources. Airlines may employ analytics to determine ticket prices. Hotels may employ Analyzing models to forecast the number of visitors on a particular night in order to maximize capacity and boost revenue. Using predictive analytics, a business is likely to perform more efficiently [9]. Analyzing is also useful for the E-commerce websites to forecast the business and understand the amount of purchases done by the customers so that they can stock up or down the products.
- **Clinical Decision Support System**: For diagnostic testing, expert systems based on predictive models may be used. It could also be employed in the development of disease-specific medications [10].

SERIES OF VARIOUS DIFFERENT PREDICTION BASED SYSTEMS :

Predicting Forecasting is the broad definition of predictive modeling, which is the grading of information using predictive models and subsequently predicting. However, in speaking, it can be used to refer to the professions associated with analysis. These fields include analysis of the data, which is employed in managerial decision - making. These fields can be classified as follows:

- **Predictive Models**: Forecasting analytics replicate the connection between both the efficiency and attributes of such a division. This metric represents the likelihood of an equivalent unit in a sample exhibiting the same effectiveness. When consumer performance answers are required, this method is often employed in advertising. It imitates human behaviour in order to respond to a certain query. During a connected client, it analyzes to determine the risk connected with the consumer or transaction.
- **Descriptive Models**: In addition to identifying consumers or categories in a possibility, the descriptive model establishes the link between both the information. Whereas future predictions identify a specific customer or achievement, descriptive models detect a large number of relationships between a business and its consumers. Rather than evaluating consumers based on their activities, it categorizes them based on the performance of their products. In descriptive modelling, a large number of specific agents can be modeled together just to create a forecast.

- **Decision Models**: The link between the facts, the choice, and the outcome of a judgement prediction is defined by the decision models. In addition to generating a prognosis on the outcome of a choice combining numerous aspects, this link is mentioned in the decision-making process. These

techniques are used to increase the likelihood of a given result, decrease the likelihood of a different outcome, and improve. It's performed in the creation of business rules to ensure that each consumer or scenario gets the action they need.
A sophisticated analytics system is considered as a standard that forecasts with a high degree of accuracy. It generates a prediction score for each individual. It's more like a machine that learns from past events to predict a person's future behaviour. This allows you to make more informed decisions. The accuracy of the model is determined by the quality of the data analysis.

PREDICTIVE ANALYTICS METHODOLOGIES :
All advanced analytical models are divided into two categories: classification models and regression models. Classification model suggests whether or not variables belong to a specific class, whereas regression models suggest a quantity. We will now look onto the main approaches which would be usually noticed utilized in the improvement of forecasting analytics.

DECİSİON TREE :
A logistic regression is a type of categorization that may also be applied to regression analysis. It's a tree-like concept that links acts to their possible consequences [11]. The consequences could be the effect of events, the cost of materials, or the utility of the substance. Each branch in the tree-like structure represents a choice among several options, and each leaf represents a choice. It separates data into subgroups depending on input variable categorization. It assists people in making decisions. Decision trees are popular because they are simple to understand and interpret. Figure 2 displays a popular decision tree model.

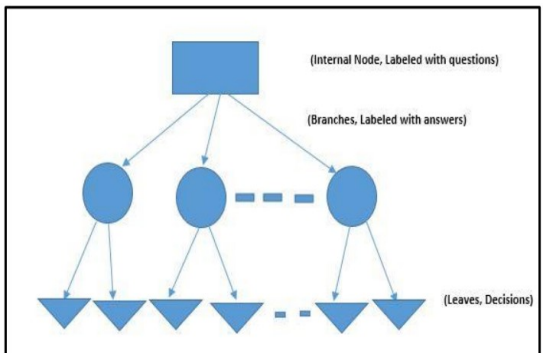

Fig. 2. Decision Tree

A decision tree is a classification algorithm that may be applied to simple regression as well. It's a tree-like model that connects acts with their possible consequences [11]. The consequences could be as a result of the experiences, material costs, or utility. Each leaf represents a decision, and each stem in the tree-like architecture represents a choice amongst a variety of options. Based on the classifications of input parameters, it separates data into subgroups. Customers can use it to help them make decisions. Decision trees are popular since they're simple to understand and comprehend. The decision tree is shown in Figure 2 as a popular model.Besides these characteristics, decision trees offer a number of pro's and downsides. New hypothetical situations may be highlighted in this part, demonstrating the model's flexibility and adaptability. It can be used with other decision models if necessary. They are constrained in their ability to implement the modifications. A minor modification in the available results in a significant change in the structure. In terms of prediction accuracy, they trail behind the other predictive algorithms. The computations in this system is complicated, providing given the usage of unknown data.

ISSN: 2278-4632
Vol-14, Issue-7, No.01,July: 2024

REGRESSİON MODEL :

Correlation is a popular statistical method for determining the connection between two or more independent variables. It models the relationship between one or more independent variables.

It examines how well the quantity of the dependent variable fluctuates when at least one of the simulated relationship's number of distinct varies. [12]. Figure 3 depicts the modelled relationship between dependent and independent relationships.

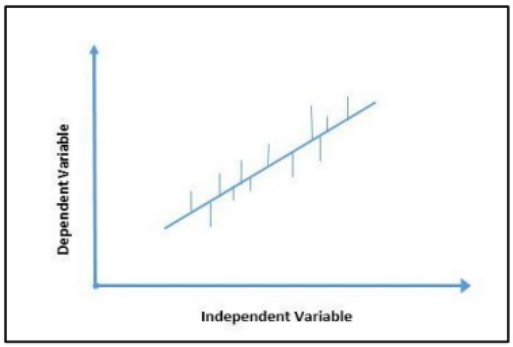

Fig. 3. Regression Model Depiction

In the case of continuous data with a normally distributed, this regression based system identifies the main trend in huge sets of data. It's being used to determine how specific factors affect the movement of a parameter. A regression model is used to estimate the price of a dependent variables in regression. In this situation, a regression function is used to all of the independent variables to map them to the response variable. The forecast of the regression model using a probabilistic model characterizes the fluctuation of the dependent variable in this approach.The two categories of regression methods adopted in predictive analysis for prediction or forecast are linear regression and logistic regression. The linear regression analysis is being used to illustrate the linear relationship between two variables. The regression function in this scenario is a linear function. When there are categories of response variables, the logistic regression can be used on either side. This model predicts unknown amounts of categorical data based on known values of the independent variables. It can only consider a limited number of factors while making predictions.

ARTİFİCİAL NEURAL NETWORK :

The human nervous system's skills of absorbing input signals and creating outputs are simulated by an artificial neural network, which is a channel of artificial neurons based on biological neurons [13]. This is a powerful model capable of simulating incredibly complicated relationships. Figure 5 depicts the design of a multipurpose artificial neural network.

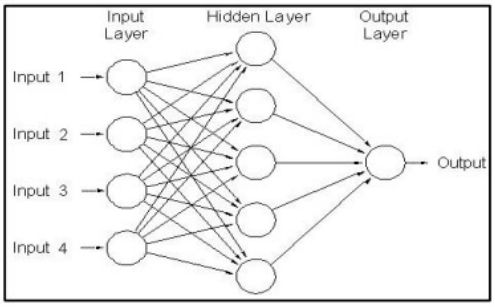

Fig. 4. ANN Depiction

Convolutional neural nets are employed in prediction-based analyses applications as a strong way of acquiring knowledge from sample databases and making predictions on latest data. An input sequence of the training examples is used for handling through the program's input nodes and is sent to the hidden nodes, which is a matrix of layers. Based on the output need, neurons utilize a variety of activation functions. The output of one neuron is passed to act as the input of the other neuron of the following layer. Output is gathered at the output layer, which might be a forecast based on fresh data.There are many types of artificial neural networks, each with its own algorithm. Backpropagation is a well-known technique that is commonly used in supervised training assignments. Convolutional layers can also help with unsupervised learning problems. Clustering, like deep neural network, is a technique used in unsupervised learning. They can work with data that has non-linear connections. They're frequently used to evaluate linear regression and decision tree results. These algorithms, which are capable of pattern classification, are utilised in picture recognition challenges.

BAYESİAN STATİSTİCS :

This approach is a type of analytics in which characteristics are treated as random numbers and the phrase "degree of belief" is used to denote the likelihood of an event occurring [14].

The Bayesian principle, essentially specifies priori and posteriori events, underpins Bayesian analysis. The goal of probabilistic reasoning is to figure out the chances of an a posteriori happening after an a priori event has occurred. The Bayes' theorem, but at the other side, determines the probability of an a priori incident happening if an a posteriori event has already occurred. Figure 5 shows an example of this.

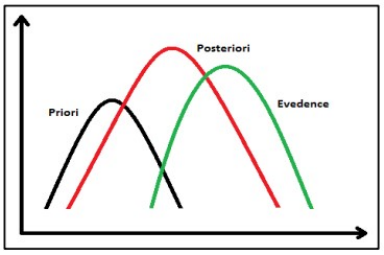

Fig. 5. Bayesian Model

It employs a probabilistic topic model called as the Bayesian network to describe the conditional relationships between random variables. This idea may be used to determine the causes based on the outcomes of those causes. For instance, it may be used to diagnose a disease based on its characteristics.

ENSEMBLE LEARNİNG :

This is classified as a supervised algorithm in the field of data analytics. Here the models are generated by learning multiple comparable types of models and then merging their prediction results. As a result, the model's accuracy improves. This method of development reduces the model's bias and variance. It aids in determining the optimal model to utilize with fresh data [15]. Figure 6 is an example of categorization using ensemble techniques.

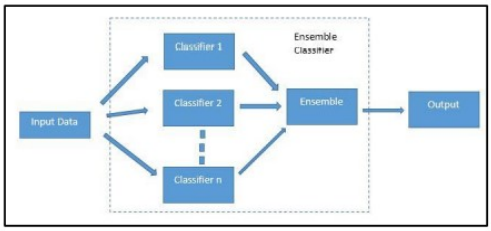

Fig. 6. Ensemble Classifier

GRADİENT BOOST MODEL :

As a machine learning technique, this technique is used in predictive modeling. It is most commonly utilized in classification or regression purposes. It is similar to an ensemble model in that it aggregates the predictions of decision trees that are poor predictive models [16]. It's a promoting approach in which the dataset is resampled multiple times and the outcomes are calculated as a weighting factor of heavily processed database. It has the advantage of being less prone to overfitting, which is a flaw in many machine learning algorithms. The usage of decision trees in this method aids in the accurate fitting of data, while enhancing improves data fitting.Figure 7 shows a representation of this paradigm.

Fig. 7. Gradient Boosting

SUPPORT VECTOR MACHİNE :

It is a frequently used supervised machine learning technique in predictive modeling. It uses association learning techniques to look at data for regression and classification [17] [18]. Nevertheless, it is mostly used in classification tasks. It's a categorization method that divides examples into groups using a hyperplane. It is the representation of instances in a plane in such a way that the examples may be clearly classified. Depending on where the new samples fall on the gap, they are then estimated to belong to a class. A separation using a support vector machine is shown in Figure 8.

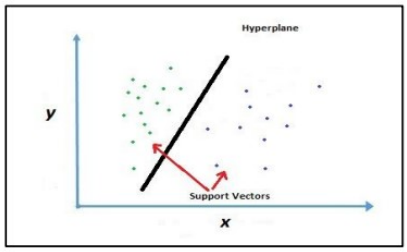

Fig. 8. Support Vector Machine

TİME SERİES ANALYSİS :

Time analysis is used to examine that employs time series data collected at regular basis over a period of time. It combines traditional data mining algorithms with forecasting [19]. The two domains of time series data are frequency analysis and time domain.

It forecasts the future of a parameter at upcoming time intervals based on an examination of historical time interval values. It is often utilised in stock market forecasting and weather forecasting. Figure 9 depicts an example of price variance over time for a product, as well as forecasted patterns for future years.

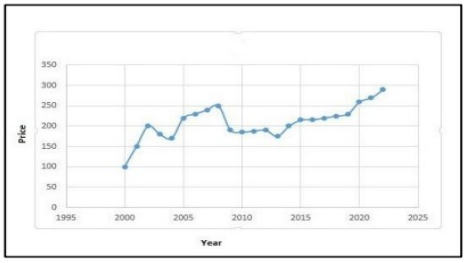

Fig. 9. Time series Model

ISSN: 2278-4632
Vol-14, Issue-7, No.01,July: 2024

k-NEAREST NEİGHBORS (k-NN) :

This is a a non-parametric way to solving classification and regression problems. This method uses the k closest training instances in a feature set as input [20]. The identification of a class is the result of a classification task, while the quantity of an item's property is the result of a forecasting model. Machine learning algorithms of this type are the most basic. Figure 10 shows an example of k-NN regression.

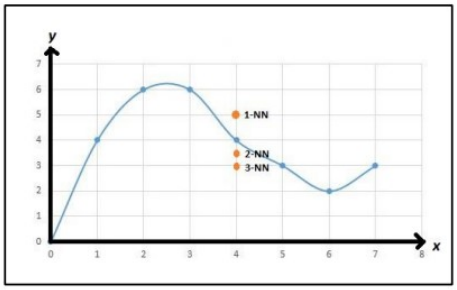

Fig. 10.KNN Regression Analysis

PRİNCİPLE COMPONENT ANALYSİS :

It is a statistical procedure that is commonly employed in prediction models for exploratory data analysis. It is deeply linked to factor analysis, that is used to solve the eigenvalues of a matrices. It is also used to describe the variance in a dataset [21]. Figure 11 depicts an example of a principal component in a dataset.

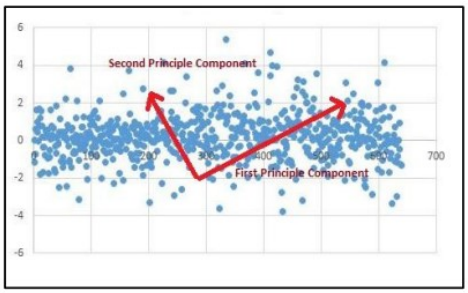

Fig. 11. Principle Component Analysis

PREDICITVE ANALYTICS USE IN REAL LIFE SCENARIOS :

Predictive analytics has various applications in a wide variety of fields. From medical decision analytics to stock market prediction, where a sickness may be identified based on indications and investment gains can be anticipated, some of the most popular uses are listed below.

BANKİNG AND FİNANCİAL SERVİCES :

Predictive analysis is widely used in the finance and banking industries. In all businesses, information and finance are critical components, and gaining insight from some of those information and the movement of money is essential. Predictive analytics assists in the detection of dishonest customers and shady transactions. It lowers the credit risk these businesses face when lending money to their customers. Cross-sell and up-sell possibilities, as well as retaining existing customers and attractiveness, are all aided by it. In the financial sector, where money is being invested in shares or other commodities, predictive modelling predicts the value for money and supports in investment decision-making.

RETAİL INDUSTRY :

Predictive modeling assists the retail industry in analyzing customer and determining whatever they want and desire. They use this method to forecast user behavior toward an item. Businesses may establish prices and special offers on items after analyzing client purchasing behavior. It also assists the retail business in projecting how a certain product will do during a given season. They may advertise their items and contact clients with offers and pricing tailored to certain customers.

HEALTH AND INSURANCE :

Quantitative methods is used in the pharma industry to improve medication supply chain management and medicinal chemistry. These corporations can use this approach to estimate the expiration of pharmaceuticals in a given location owing to a lack of sales. Customers' fraud claims are identified and predicted using predictive analytics models in the insurance industry. The health insurance industry employs this technology to identify consumers where they are having high chances of imperiling of developing a major ailment and approach them with the best insurance policies for their expenditure [23].

OİL GAS AND UTİLİTİES :

The fossil fuel industries utilise prediction data to assess equipment breakdown in order to reduce risk. These simulations are used to anticipate future resource needs. Energy-based businesses can anticipate the need for maintenance in order to avoid a fatal accident in the coming years [24].

GOVERNMENT AND PUBLİC SECTOR :

Government entities employ big data-based predictive analytics tools to identify probable criminal behaviour in a given region. They use social media data to learn about suspects' past behaviour and predict their future behaviour. Predictive modelling is being used by organizations to forecast population patterns at the national and state levels. Data analysis techniques are frequently used to improve cybersecurity [25].

CONCLUSION AND FUTURE SCOPE :

The use of estimation techniques in probabilistic models has a long history. Previously, statistical models were used as predictive models, but they were based on a small sample of data from a large data set. Novel strategies have been introduced as computer sciences and computer simulations have progressed, and better and better programmes have been offered over time. Machine learning and artificial intelligence breakthroughs have changed the technology landscape by offering intelligent computation methodologies and algorithms.Convolutional neural networks ushered in a revolution in the field of predictive analysis. Based on the input parameters, any valuation outcome or future can be predicted. There is now a trend of using deep learning models in data analytics, but they are being used extensively in this activity, thanks to advancements in the field of machine learning and the introduction of deep learning models. This study suggests that various algorithms for predictive analytics could be developed as a result of this study. There's also the potential of adding new features to existing models in order to improve their work performance.

REFERENCES :

1. Charles Elkan: Predictive analytics and data mining, University of California, San Diego.(2013)
2. Eric Siegel: Predictive Analytics, John Willey and Sons Ltd. (2016)
3. Charles Nyce: Predictive Analytics White Paper, American Institute of CPCU/IIA(2007).
4. W Eckerson: Extending the Value of Your Data Warehousing Investment, The Data Warehouse Institute (2007)

5. Sue Korn: The Opportunity of Predictive Analytics in Finance, HPC Wire. (2011)
6. M Nigrini: Forensic Analytics: Methods and Techniques for Forensic Accounting Investigations, John Willey and Sons Ltd. (2011)
7. M Schiff : BI Experts: Why Predictive Analytics Will Continue to Grow, The Data Warehouse Institute (2012)
8. F Reichheld, P Schefter, Retrieved: The Economics of E-Loyalty, Harvard Business School Working Knowledge (2018)
9. V Dhar: Predictions in Financial Markets: The Case of Small Disjuncts, ACM Transaction on Intelligent Systems and Technology, Vol-2, Issue-3 (2001)
10. J Osheroff, J Teich, B Middleton, E Steen, A Wright, D Detmer: A Roadmap for National Action on Clinical Decision Support, JAMIA: A Scholarly Journal of Informatics in Health and Biomedicine, Vol-14, Issue2, Pages-141-145(2007)
11. B Kaminski, M Jakubczyk, P Szufel: A framework for sensitivity analysis of decision trees, Central European Journal of Operations Research, Vol26, Issue-1, Pages-135-159 (2018)
12. J S Armstrong: Illusions in regression analysis, International Journal of Forecasting, Vol-28, Issue-3, Pages-689-694 (2012)
13. W S McCulloch, Walter Pitts: A logical calculus of the ideas immanent in nervous activities, The bulletin of mathematical biophysics, Vol-5, Issue-4, Pages-115- 133(1943)
14. Peter M Lee: Bayesian Statistics: An Introduction, 4th Edition, John Willey and Sons Ltd (2012)
15. R Polikar: Ensemble based Systems in decision making, IEEE Circuits and Systems Magazine, Vol-6, Issue-3, Pages-21-45 (2006)
16. J H Friedman: Greedy Function Approximation: A Gradient Boosting Machine, Lecture notes (1999)
17. C Cortes : Support-vector networks, Machine Learning, Vol-20, Issue-3, Pages- 273-297 (1995)
18. Ben Hur et al: Support Vector Clustering, Journal of Machine Learning Research, Vol-2, Pages125-137 (2001)
19. J Lin, E Keogh, S Lonardi, C Chiu: 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, (2003)
20. N S Altman: An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression, The American Statistician, Vol-46, Issue-3, Pages- 175-185 (1992)
21. H Abdi, L J Williams: Principal component analysis, WIREs: Computational Statistics, Vol-2, Issue-4, Pages-433-459 (2010)
22. K Das, GS Vidyashankar: Competitive Advantage in Retail Through Analytics: Developing Insights, Creating Values, Information Management (2006)
23. N Conz: Insurers Shift to Customer-Focused Predictive Analytics Technologies, Insurance & Technology, (2008)
24. J Feblowitz :Analytics in Oil and Gas: The Big Deal About Big Data, Proceeding of SPE Digital Energy Conference, Texas, USA, (2013)
25. G H Kim, S Trimi, J-H Chung: Big-data applications in the government sector, Communications of the ACM, Vol-57, Issue-3, Pages-78, (2014).

