EXCEL ENGINEERING COLLEGE
(Autonomous)
B.E Computer Science and Engineering
V Semester

20CSE04- BUSINESS INTELLIGENCE

Regulations 2020

Question Bank with Answer Key

UNIT – III (**BUSINESS FORECASTING**)

PART- A

Q.No	Questions
1	**Define Predictive Modeling.** • Predictive modeling refers to the process of using statistical algorithms and machine learning techniques to create models that can make predictions about future outcomes based on historical data. • The goal of predictive modeling is to identify patterns, relationships, and trends within a dataset to make informed predictions about unknown or future events.
2	**What is data mining?** • Data mining is the process of discovering patterns, trends, correlations, or meaningful insights from large sets of data. It involves the use of various techniques and algorithms to analyze and extract valuable information from raw data. • The goal of data mining is to uncover hidden patterns and knowledge that can be used for decision-making, prediction, and optimization.
3	**Write down the steps for predictive analysis using Machine Learning.** • Step 1: Define the problem statement. • Step 2: Collect the data. • Step 3: Clean the data. • Step 4: Perform Exploratory Data Analysis (EDA). • Step 5: Build a predictive model. • Step 6: Validate the model. • Step 7: Deploy the model. • Step 8: Monitor the model.
4	**List the types of Business Forecasting.** • Business forecasting involves making informed guesses about certain business metrics, regardless of whether they reflect the specifics of a business, such as sales growth, or predictions for the economy as a whole. • Financial and operational decisions are made based on economic conditions and how the future looks, albeit uncertain.
5	**Define Exploratory Data Analysis.** • Exploratory Data Analysis (EDA) is an approach to analyzing and visualizing data sets to summarize their main characteristics, often with the help of graphical representations. • The primary goal of EDA is to understand the underlying structure, patterns, and relationships within the data, leading to insights that can inform further analysis or guide decision-making.

6	**List the approaches of Quantitative model.** • Quantitative models are mathematical representations of real-world phenomena used to analyze and make predictions based on quantitative data. • There are various approaches to quantitative modeling, each with its specific characteristics and applications.
7	**How is BI process used in Private and Public Intelligence?** • As part of the BI process, organizations collect data from internal IT systems and external sources, prepare it for analysis, run queries against the data and create data visualizations. • BI dashboards and reports to make the analytics results available to business users for operational decision-making and strategic process.
8	**What is data reduction?** • Data reduction refers to the process of reducing the volume but producing the same or similar analytical results by transforming or aggregating raw data into a more compact, summarized form. • This is often done to handle large datasets more efficiently, decrease processing time, and simplify the analysis without losing essential information. Several techniques are commonly employed for data reduction.
9	**Name the importance of business intelligence system.** • The purpose of BI is to help inform and improve business decision-making by making data easier to interpret and act on. • While BI involves collecting and visualizing data, the term also refers to the software tools that carry out these practices.
10	**Differentiate between Quantitative model and Qualitative model.** • Quantitative models primarily deal with numerical and measurable data. This includes variables with a clear numerical representation, such as sales figures, revenue, quantities, or time. • Qualitative models deal with non-numeric, categorical, or subjective data. This can include textual information, opinions, perceptions, or descriptive data that may not have a direct numerical representation.
11	**List the approaches of Quantitative model.** • Quantitative models involve the use of numerical data and mathematical methods to analyze and make predictions. • There are various approaches within quantitative modeling, each suited to different types of problems and objectives.
12	**Differentiate market research and Delphi model.** • Market Research: Involves the systematic gathering, analysis, and interpretation of data related to a specific market, including information about consumers, competitors, and the overall industry. • Delphi Model: A structured and iterative forecasting method that relies on the input of a panel of experts or stakeholders. It is used to reach a consensus on a complex or uncertain issue.
13	**Why is predictive analysis important?** • Predictive analytics is the use of data, statistical algorithms and machine learning techniques to identify the likelihood of future outcomes based on historical data. • The goal is to go beyond knowing what has happened to providing a best assessment of what will happen in the future.
14	**Give the Uses Cases For Machine Learning Based Predictive Analytics.**

	• Predictive analytics and machine learning go hand-in-hand, as predictive models typically include a machine learning algorithm. • These models can be trained over time to respond to new data or values, delivering the results the business needs. Predictive modelling largely overlaps with the field of machine learning.
15	**How does predictive analytics work?** • Predictive analytics works by leveraging statistical algorithms, machine learning models, and data mining techniques to analyze historical data, identify patterns, and make predictions about future events or outcomes. • Predictive analytics is a branch of advanced analytics that uses data, statistical algorithms, and machine learning techniques to identify the likelihood of future outcomes based on historical data.
16	**Write the most widely used predictive models.** • Decision trees: Decision trees are a simple, but powerful form of multiple variable analysis. • Regression (linear and logistic) Regression is one of the most popular methods in statistics. • Neural networks.
17	**List some Applications of predictive analytics and machine learning.** • Financial Services. Reduce risk by automatically analyzing credit risk or loan default likelihood. • Healthcare. Provide better patient care based on patient admission and readmission forecasting.
18	**Differentiate Forecasting vs predictive analytics: which is more accurate?** • Forecasting involves estimating future values based on historical data and patterns. It often relies on quantitative methods to project trends and make predictions. • Predictive analytics is a broader term that encompasses various statistical and machine learning techniques to analyze current and historical data and make predictions about future events or trends.
19	**Differentiate Business Intelligence and Data Mining.** • Business Intelligence (BI): BI involves the use of technologies, processes, and tools to gather, analyze, and present business information. It focuses on providing historical, current, and predictive views of business operations to support decision-making. • Data Mining: Data mining is a specific process within the broader field of data analysis. It involves discovering patterns, correlations, and trends in large datasets to extract valuable information and make predictions.
20	**Write the Challenges of predictive analysis in data mining.** • The accuracy of predictive analytics models is limited by the completeness and accuracy of the data being used. • The analytical algorithms attempt to build models based on the available data, deficiencies in the data may lead to deficiencies in the model.

PART- B

Q. No	Questions
1	**Discuss in detail how machine learning improves predictive analysis in business intelligence.** • Machine learning (ML) plays a crucial role in enhancing predictive analysis within the realm of business intelligence (BI). By leveraging advanced algorithms and statistical models, machine learning contributes to more accurate predictions, deeper insights, and improved decision-making. Here's a detailed discussion on how machine learning improves predictive analysis in business intelligence: **Improved Accuracy and Precision:** • Machine learning algorithms can analyze large and complex datasets, identifying patterns and relationships that may be challenging for traditional statistical methods. This leads to more accurate and precise predictions, enabling businesses to make informed decisions based on reliable insights. **Complex Pattern Recognition:** • Machine learning excels in recognizing intricate patterns and non-linear relationships within data. This is particularly beneficial when dealing with multifaceted business scenarios where traditional BI tools may struggle to capture complex interactions among variables. **Predictive Modeling for Forecasting:** • ML algorithms, such as regression models, decision trees, and time series models, are used for forecasting future trends and outcomes. This is invaluable for businesses in predicting sales, demand, customer behavior, and other critical factors influencing performance. **Customer Segmentation and Personalization:** • ML algorithms enable businesses to segment their customer base more effectively. This segmentation can be used for personalized marketing, product recommendations, and tailoring services to specific customer needs, thereby improving customer satisfaction and retention. **Fraud Detection and Risk Management:** • ML models, particularly anomaly detection algorithms, enhance the ability to detect irregularities and potential fraud patterns. In finance and other industries, this is crucial for identifying fraudulent activities and managing risks effectively. **Churn Prediction and Customer Retention:** • Machine learning algorithms can predict customer churn by analyzing historical data and identifying factors contributing to customer attrition. This allows businesses to proactively implement strategies for customer retention, such as targeted promotions or personalized engagement. **Supply Chain Optimization:** • ML models contribute to optimizing supply chain operations by predicting demand, identifying potential bottlenecks, and streamlining inventory management. This helps businesses improve efficiency, reduce costs, and enhance overall supply chain performance.

	Sentiment Analysis: • Natural Language Processing (NLP) techniques, a subset of machine learning, are employed for sentiment analysis of textual data. This is valuable for understanding customer opinions, social media trends, and market sentiments, providing businesses with insights into public perception.
2	**Elaborate in detail about business forecasting and planning techniques.** • Business forecasting and planning are essential processes that enable organizations to anticipate future trends, make informed decisions, and set strategic goals. These techniques involve analyzing historical data, identifying patterns, and using various methods to predict future outcomes. Here are some key business forecasting and planning techniques: **Time Series Analysis:** • Definition: Time series analysis involves studying past data points to predict future values based on the assumption that future trends will follow historical patterns. • Methods: Moving averages, exponential smoothing, and autoregressive integrated moving average (ARIMA) are common techniques in time series analysis. **Regression Analysis:** • Definition: Regression analysis examines the relationship between a dependent variable and one or more independent variables. It is used to predict the value of the dependent variable based on the values of the independent variables. • Methods: Linear regression, multiple regression, and logistic regression are common types of regression analysis. **Scenario Analysis:** • Definition: Scenario analysis involves creating multiple hypothetical scenarios to assess the impact of different factors on business outcomes. It helps identify potential risks and opportunities. • Methods: Constructing best-case, worst-case, and base-case scenarios to understand the range of possible outcomes. **Market Research and Surveys:** • Definition: Gathering data through market research and surveys to understand customer preferences, market trends, and industry dynamics. • Methods: Conducting customer surveys, focus groups, and analyzing market reports to gather qualitative and quantitative data. **Delphi Method:** • Definition: The Delphi method involves obtaining expert opinions through a structured and iterative process. Experts provide forecasts anonymously, and their feedback is used to refine predictions. • Methods: Iterative rounds of questionnaires and feedback until a consensus or convergence of opinions is reached. **Simulation Models:** • Definition: Simulation models use computer programs to imitate the behavior of a system over time. They allow businesses to assess the impact of different variables on outcomes. • Methods: Monte Carlo simulations and system dynamics modeling are examples of simulation techniques. **Econometric Models:**

	• Definition: Econometric models apply statistical methods to economic data to analyze and forecast economic trends. They help businesses understand the relationship between economic factors and their impact on business operations. • Methods: VAR (Vector Autoregression) models and simultaneous equation models are examples of econometric modeling. **Neural Networks and Machine Learning:** • Definition: Utilizing advanced machine learning algorithms, such as neural networks, to analyze patterns in large datasets and make predictions. • Methods: Artificial neural networks, decision trees, and ensemble methods like Random Forest or Gradient Boosting. **Forecasting with Leading Indicators:** • Definition: Identifying leading indicators that precede changes in business conditions. These indicators can provide early signals of potential trends. • Methods: Monitoring economic indicators, consumer confidence indices, and other metrics that have historically shown a correlation with future business performance. **SWOT Analysis:** • Definition: A SWOT analysis evaluates an organization's strengths, weaknesses, opportunities, and threats. It helps in strategic planning by identifying internal and external factors that can impact business outcomes. • Methods: Identifying strengths and weaknesses (internal factors) and opportunities and threats (external factors) to formulate strategies based on the analysis. **Trend Extrapolation:** • Definition: Extrapolating existing trends into the future based on historical data. This method assumes that past trends will continue.
3	**Differentiate Data mining and Predictive analysis. Write the advantages, Strategies of Predictive analytics.** • **Data Mining**: Aims to uncover hidden patterns and relationships within data, which can be useful for understanding trends, segmenting data, and gaining insights into the underlying structure of the information. • Methods: • **Data Mining**: Involves unsupervised learning techniques, where the algorithms explore the data without predefined outcomes or targets. • Applications: • **Data Mining**: Used for knowledge discovery, clustering similar data points, finding associations between variables, and identifying outliers or anomalies. • Predictive Analytics: • Definition: • **Predictive Analytics**: Predictive analytics is the process of using statistical algorithms and machine learning techniques to analyze historical and current data in order to make predictions about future events or trends. • Focus: • Predictive Analytics: Primarily focuses on making predictions and forecasting future outcomes based on patterns identified in historical data. • Objective: • **Predictive Analytics**: Aims to generate actionable insights and predictions that can inform decision-making and planning for future events. • Methods:

	• **Predictive Analytics**: Involves supervised learning techniques, where the algorithms are trained on historical data with known outcomes to make predictions on new, unseen data. • Applications: • **Predictive Analytics**: Used for forecasting sales, predicting customer behavior, identifying potential fraud, optimizing business processes, and making data-driven decisions. • Advantages of Predictive Analytics: • **Informed Decision-Making**: • Predictive analytics provides organizations with insights and predictions, enabling informed decision-making and strategic planning. • **Risk Mitigation:** • By identifying potential risks and uncertainties, predictive analytics helps organizations proactively manage and mitigate risks, reducing the likelihood of negative outcomes. • **Resource Optimization**: • Businesses can optimize resource allocation, whether it's in terms of inventory, workforce, or marketing spend, based on predicted future demands and trends. • Enhanced Customer Experience: • Predictive analytics enables personalized and targeted marketing efforts, leading to improved customer satisfaction and retention. • **Fraud Detection:** • In financial services and other industries, predictive analytics is effective in detecting anomalous patterns indicative of fraud, helping prevent financial losses. • **Operational Efficiency:** • Organizations can optimize operational processes by predicting equipment failures, identifying bottlenecks, and streamlining workflows, leading to improved efficiency. • Competitive Advantage: • By leveraging predictive analytics, businesses can gain a competitive edge by anticipating market trends and customer preferences, enabling them to respond quickly to changing conditions.
4	**Explain in detail about steps for Predictive analysis using Machine Learning.** • Predictive analysis using machine learning involves a systematic process to build, train, and deploy models that can make predictions on new, unseen data. Here is a detailed explanation of the steps involved in predictive analysis using machine learning: **Define the Problem:** • Clearly define the problem or question you want to address with predictive analysis. Understand the specific business goals and outcomes you aim to achieve through predictions. This step is crucial for guiding the entire predictive modeling process. **Data Collection:** • Gather relevant data from various sources. This may include historical records, customer information, transactional data, and any other data that is pertinent to the problem at hand. Ensure the data is representative and covers the necessary time periods. Acquiring high-quality and diverse data is essential for building effective predictive models. **Data Cleaning and Preprocessing:**

	• Clean and preprocess the data to ensure it's accurate, consistent, and ready for analysis. This involves handling missing values, removing outliers, and transforming variables. The goal is to prepare the data for machine learning algorithms by addressing issues that may negatively impact model performance. **Exploratory Data Analysis (EDA):** • Conduct exploratory data analysis to understand the characteristics of the data. Visualize distributions, identify patterns, and explore relationships between variables. EDA helps in gaining insights that can inform feature selection and preprocessing decisions. **Feature Selection and Engineering:** • Identify and select relevant features (variables) that will be used to make predictions. Additionally, create new features through feature engineering if necessary. Feature selection is crucial for building models that are both accurate and efficient. **Data Splitting:** • Split the dataset into training and testing sets. The training set is used to train the machine learning model, while the testing set is reserved to evaluate the model's performance on unseen data. Common splits include 80% for training and 20% for testing, but this can vary based on the dataset size and specific requirements. **Model Selection:** • Choose the appropriate machine learning model based on the nature of the problem. Consider factors such as the type of prediction (classification or regression), the complexity of relationships, and the characteristics of the data. Common models include linear regression, decision trees, support vector machines, and neural networks. **Model Training:** • Train the selected model using the training data. The model learns patterns and relationships within the data to make predictions. The training process involves adjusting the model's parameters to minimize the difference between its predictions and the actual outcomes. **Validation and Hyper parameter Tuning:** • Validate the model's performance using the testing set. Adjust hyperparameters, which are configuration settings that control the learning process, to optimize the model's performance. Techniques such as cross-validation may be employed to assess how well the model generalizes to new data. **Evaluation Metrics:** • Choose appropriate evaluation metrics based on the type of prediction. For classification problems, metrics such as accuracy, precision, recall, and F1 score are commonly used. For regression problems, metrics like mean squared error (MSE) or root mean squared error (RMSE) are typical.
5	**Explain in detail about types of data mining model with neat sketch.** • Data mining models are mathematical algorithms or computational techniques that are used to discover patterns, trends, associations, and knowledge from large datasets. These models help in extracting valuable information and making predictions or decisions based on the analyzed data. • The motive of data mining is to recognize valid, probable advantageous, and understandable connections and patterns in existing data. Database technology has become more developed where huge amounts of data require to be stored in a

database, and the wealth of knowledge hidden in those datasets is collected by business people as a usable tool for making business vital decisions.

- Data mining then Fascinate more awareness as it obligated to take out valuable information from the raw data that businesses can use to enlarge their advantageously via a profitable decision-making process.
- Data mining is used to depict intelligence in databases; it is a procedure of extracting and recognize useful information and succeeding knowledge from databases using mathematical, statistical, artificial intelligence, and machine learning technique.
- Data mining consolidates many various algorithms to put through different tasks. All these algorithms assimilate the model into the data. The algorithms examine the data and modulate the data that is closest to the features of the data being examined. Data mining algorithms can be described as consisting of three parts.
- Model – The objective of the model is to fit the model in the data.
- Preference – Some identification tests must be used to fit one model over another.
- Search – All algorithms are necessary for processing to find data.

Types of Data Mining Models

PredictiveModel:

A predictive model constitutes prediction concern values of data using known results found from various data. Predictive modelling may be made based on the use of variant historical data. Predictive model data mining tasks comprise regression, time series analysis,classificationprediction.

The Predictive Model is known as **Statistical Regression**. It is a monitoring learning technique that Incorporates an explication of the dependency of few attribute values upon the values of other attributes In a similar item and the growth of a model that can predict these attribute values for recent cases.

- **Classification**
  It is the act of assigning objects to one of several predefined categories. Or we can define classification as a learning function of a target function that sets each attribute to a predefined class label.
- **Regression**
  It is used for appropriate data. It is a technique that verifies data values for a function. There are two types of regression –
  1. **Linear Regression** is associated with the search for the optimal line to fit the two attributes so that one attribute can be applied to predict the other.
  2. **Multi-Linear Regression** involves two or more than two attributes and data are fit to multidimensional space.
- **TimeSeriesAnalysis**
  It is a set of data based on time. Time series analysis serves as an independent variable to estimate the dependent variable in time.
- **Prediction**
  It predicts some missing or unknown values.

DescriptionModel :

A descriptive model distinguishes relationships or patterns in data. Unlike Predictive Model, a descriptive model serves as a way to explore the properties of data being examined, not to predict new properties, clustering, summarization, associating rules, and sequence discovery are descriptive model data mining tasks.

Descriptive analytics Concentrate on the summarization and conversion of the data into significant information for monitoring and reporting.

- **Clustering**
  It is the technique of converting a group of abstract objects into classes of identical objects.

	• **Summarization** It holds a set of data in a more in-depth, easy-to-understand form. • **AssociativeRules** They find an exciting consistency or causal relationship between a large set of data objects. • **Sequence –** It is the discovery of interesting patterns in the data is in relation to some objective or subjective measurement of how interesting it is.
6	**Differentiate the analysis of Logic-Based Models and Data-Driven Models in Decision Making.** • Logic-based models and data-driven models represent two distinct approaches to decision making, each with its own set of characteristics and methodologies. Here's a differentiation between the analysis of logic-based models and data-driven models in decision making: **Logic-Based Models:** **Basis:** • Logic-Based Models: Rely on explicit rules, knowledge, and logical reasoning. Decision-making is guided by predefined rules and relationships. • Data-Driven Models: Leverage patterns and insights derived directly from data. Decision-making is based on observed trends and relationships in the data. **Rule-based Reasoning:** • Logic-Based Models: Use if-then rules, deductive reasoning, and expert knowledge to make decisions. • Data-Driven Models: Discover patterns and correlations in the data, allowing for inductive reasoning and predictive modeling. **Knowledge Representation:** • Logic-Based Models: Utilize explicit knowledge representation, often in the form of rules or a knowledge base. • Data-Driven Models: Extract implicit knowledge from data through statistical methods, machine learning algorithms, or neural networks. **Domain Expertise:** • Logic-Based Models: Require significant input from domain experts to define rules and logical relationships. • Data-Driven Models: Can automatically learn patterns and relationships from data, reducing the need for explicit domain knowledge. **Flexibility:** • Logic-Based Models: May be less flexible in adapting to changing or complex environments, as the rules need to be explicitly defined. **Data-Driven Models**: Exhibit flexibility in handling complex and dynamic datasets, adapting to changes and evolving patterns. **Data-Driven Models:** **Dependency on Data:** • Logic-Based Models: Less dependent on large volumes of historical data and can operate with a limited dataset. **Data-Driven Models:** Require substantial amounts of data for training and validation to identify patterns and make accurate predictions. • **Prediction and Generalization:** • Logic-Based Models: May struggle with making accurate predictions in situations where rules are not explicitly defined. • **Data-Driven Models:** • Excel in making predictions and generalizing patterns based on historical data. • **Exploratory Nature:**

	• Logic-Based Models: Tend to be more deterministic and less exploratory in nature, as decisions are based on predefined rules. • **Data-Driven Models:** • Allow for exploration of hidden patterns and insights in the data, discovering relationships that may not be apparent through explicit rules. **Scalability:** • Logic-Based Models: Can be scalable in terms of rule-based systems but might become complex and hard to manage with an increasing number of rules. • Data-Driven Models: Often scalable and can handle large datasets, but computational complexity may increase with more complex algorithms.

(Note:*Blooms Level (R – Remember, U – Understand, AP – Apply, AZ – Analyze, E – Evaluate, C – Create)
PART A- Blooms Level: Remember, Understand, Apply
PART B- Blooms Level: Understand, Apply, Analyze, Evaluate (if possible)
Marks: 16 Marks, 8+8 Marks, 10+6 Marks)

Subject In charge (Name & Signature) Course Coordinator (Name & Signature) HOD IQAC