



CLASS - VIII COMPUTER SCIENCE NOTES JANUARY

CHAPTER 11 - MACHINE LEARNING AND DATA SCIENCE

A. Tick the correct option.

1. The ability of _____ to work so well with data analytics is the main reason AI and big data are now seemingly inseparable.
(a) Computer (b) Machine learning **(c) AI**
2. Data science professionals apply machine learning algorithms to:
(a) Numbers (b) Text **(c) Both (a) and (b)**
3. _____ is a domain of artificial intelligence (AI).
(a) Physics **(b) Machine learning** (c) Computer science
4. Machine learning (ML) can be classified into _____ major learning algorithms.
(a) 4 **(b) 3** (c) 5
5. _____ is data in large volumes.
(a) Deep learning (b) Supervised learning **(c) AI**
6. Data science life cycle focuses only on the:
(a) Data (b) Modelling **(c) Both (a) and (b)**
7. _____ is a branch of applied mathematics that involves the collection, description, analysis, and inference of conclusions from quantitative data.
(a) Statistics (b) Cognitive science (c) Data science
8. _____ is an essential part of running a successful business.
(a) Statistics **(b) Data analysis** (c) Modelling
9. This falls somewhere between supervised and unsupervised learning:
(a) Semi-supervised learning
(b) Reinforcement learning
(c) None of these
10. Trial and error research and delayed reward are the most relevant features of:
(a) Semi-supervised learning

(b) Reinforcement learning

(c) Supervised learning

B. Fill in the blanks:-

1. Similar to the functions of human body parts, **Artificial Intelligence** has many domains that are dedicated to executing a particular task.
2. **Maching learning** represents the abiity of the AI algorithm to learn patterns from data and recognize them in unpublished data.
3. **Supervised learning** can apply what has been learned in the past to new data using labelled examples to predict future events.
4. **UnSupervised learning** studies how systems can infer a function to describe a hidden structure from unlabelled data.
5. **Reinforcement learning** is a method that interacts with its environment by producing actions and discovering errors or rewrd.
6. **Data Science** is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and information.

C. Write 'T' for true and 'F' for false for the following statements.

1. It is rare to see any artificial intelligence applications around us. - F
2. Domains of artificial intelligence are mutually interconnected. - F
3. Machine learning allows systems to learn and improve automatically from experience without being explicity programmed. - T
4. There is no importance of data science, artificial intelligence and machine learning in industries nowadays. - F
5. Machine learning is not interconnected with other areas of AI -F
6. Supervised learning systems can provide targets for any new entry after sufficient training -T
7. Unsupervised learning system does not find the correct exit, but it explores the data and can draw inferences from data sets to describe hidden structures from unlabelled data. - T
8. A data science life cycle is an iterative set of steps taken to deliver a data science project or product. _ T
9. There is very little career scope in data science, - F
10. Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and information from structures and unstructured data - T

D. Short answer question.

1. How AI is used in language translation?

User input spoken into a microphone is transmitted to the system sound card. The converter transforms the analogue signal into an equivalent digital signal for speech processing. The database is used to compare sound patterns to recognize words.

2. What are the major domains of artificial intelligence?

- Machine learning
- Data science
- Natural language processing
- Computer vision
- Cognitive computing
- Big data
- Deep learning

3. What is machine learning (ML)?

Machine learning is a domain of artificial intelligence (AI) that allows systems to learn and improve automatically from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn on their own.

4. Define data?

Data is the quantities characters or symbols on which operations are carried out by a computer, which can be stored and transmitted in the form of electrical signals and recorded on magnetic, optical or mechanical recording media.

5. What is big data?

Big data is a collection of massive and complex data sets and volumes that include massive amounts of data, data management capabilities, social media analytics and real-time data.

6. What is data science?

Data science is an interdisciplinary field uses scientific methods, processes, algorithms and systems to extract knowledge and information from structured and unstructured data and apply actionable knowledge and information from data to a wide range of application areas.

E. Long answer questions.

1. Briefly explain the algorithms of machine learning (ML).

Depending on the data we have and the strategy(algorithm) we use to recognize a pattern in the data, machine learning can be classified into:

- Supervised learning
- Unsupervised learning
- Semi-supervised learning

- Reinforcement learning
-

Supervised learning :

This can apply what has been learned in the past to new data labelled examples to predict future events. From the analysis of a known training data set, the training algorithm produces an inferred functions to make predictions about the output values.

Unsupervised learnin:

This is used when the information used for traing is not classified or labelled. Unsupervised learning studies how systems can infer a fnction to describe a hidden structure from unlabelled daa.

Semi-Supervised learning:

This falls somewhere between supervised and unsupervised learning, as they use both labelled and unlabelled data and a large amount of unlabelled data.

Reinforcement learning:

A learning method that interacts with its environment by producing actions and discovering errors or rewards. Trial and error research and delayed reward are the most relevant features of reinforcement learning.

2. Write down the advantages and disadvantages of machine learning (ML).

Advantages of Machine learning:

- It's a automatic
- It can handle varieties of data
- It can identify trends and patterns
- It is widely used in various fields.
- There is always a scope for advancement.
- It s considered best for education.

Disadvantages of Machine Learning:

- The chance of error or fault is more
- Data requirement is more
- It is time-consuming and requires more resources.
- There is a scope for inaccuracy in the interpretation of data
- More data is required for interpretation more space is required to store the data.

3. What are the major applications of machine learning?

Personalized Shopping:

The reason for their knowledge is the large amount of data they collect from you and advanced AI models that predict the most likely products you will buy. When you search for and find a product that you want to order bext, amazon collects everything in the background.

Finance:

AI is also commonly used in this field. Brokers actively rely on artificial intelligence algorithms to help them decide what to do with certain stocks. Unlike humans, a machine can process a large amount of information in a fraction of a second.

Social Media:

Wherever you are and whatever you do on the internet, each of your steps is recorded (click, search, like comment, etc). When you search on Google, like on facebook, or follow it on instagram, everything is recorded. When you have data, you can do it all. In that case, using all of the data stored by your tech companies lie google, facebook, and amazon can predict what you would like most from their platforms.

4. Explain the data science life cycle.

A data science life cycle is an iterative set of steps taken to deliver a data science project or product. Because every data science project and team are different, each specific data science life cycle. Some data science life cycles focus only on the data, modelling, and evaluation stages. Others are more comprehensive and start with understanding the business and end with implementation. The one you are going to learn is even more extended to include operations.

The data science life cycle consists of the following stages:

- Problem definition
- Data investigation and cleaning
- Minimum viable model
- Development and enhancements
- Data science operations

5. What is data analysis? Explain briefly and also write its types.

Data analysis is an essential part of running a successful business. When data is used effectively, it leads to a better understanding of a company's past performance and better decision-making for its future activities. There are many ways to use data, at all levels of a business's operations

For types of data analysis are used in all industries. Although we separate them into categories, they are all related to and based on each other. As you start to move from a simpler type of analysis to a more complex type of analysis, the degree of difficulty and resources required increase. At the same time, the level of knowledge and added value is also increasing.

The four types of data analysis are:

- Descriptive analysis
- Diagnostic analysis
- Predictive analysis
- Prescriptive analysis

6. Explain some of the widely used real-world applications of data science.

Some of the widely used real-world applications of data science are as follow:

Search Engines : The search engine is one of the most useful applications of data science. When we want to search for something on the internet, we mostly use search engines like Google, yahoo, bing, etc. Data science is used to fast track searches by providing suggestions and recommendations.

Transport: Data science plays a crucial role in the development of autonomous vehicles. Data science apps have even tackled traffic, with route optimization models that capture typical rush hours and weekend breaks.

Finance: Data science plays a key role in the financial sector. Financial institutions always have an issue of fraud and risk of losses. This necessitates the automation of loss analysis to carry out strategic decisions for the company.

E-Commerce: Data science plays a key role in the growth of the E-commerce sector. E-commerce platforms like Amazon, eBay, Flipkart, and Myntra use data science to make a better user experience with personalized recommendation.

Healthcare: Data science acts as a boon in the healthcare industry. It is used in tumour detection, drug discoveries, medical image analysis, genetics and genomics, predictive diagnosis modelling, and virtual medical bots.