

Data Science & Machine Learning with Python

Assignment: 01

Prepared By: Md. Ashikur Rahman (Batch-15)

Q1. What is Data Science?

According to IBM, Data science combines math and statistics, specialized programming, advanced analytics, artificial intelligence (AI), and machine learning with specific subject matter expertise to uncover actionable insights hidden in an organization's data. These insights can be used to guide decision making and strategic planning.



Q2. What is Artificial Intelligence?

Artificial Intelligence is a method of making a computer, a computer-controlled robot, or a software think intelligently like the human mind. AI is accomplished by studying the patterns of the human brain and by analyzing the cognitive process. The outcome of these studies develops intelligent software and systems.

Q3. What is Machine Learning?

Machine learning is an application of artificial intelligence that uses statistical techniques to enable computers to learn and make decisions without being explicitly programmed. It is predicated on the notion that computers can learn from data, spot patterns, and make judgments with little assistance from humans.

It is a subset of Artificial Intelligence. It is the study of making machines more human-like in their behavior and decisions by giving them the ability to learn and develop their own programs.

Q4. Future of Data Science and Real-life applications.

Data has applications in almost every field. The Data Science future is studded with career opportunities. Future of Data Science 2030 is estimated to bring opportunities in various areas of banking, finance, insurance, entertainment, telecommunication, automobile, etc. A data scientist will help grow an organization by assisting them in making better decisions.

Real life applications.

- **Healthcare:** Data science can identify and predict disease, and personalize healthcare recommendations.
- **Transportation:** Data science can optimize shipping routes in real-time.
- **Sports:** Data science can accurately evaluate athletes' performance.
- **Government:** Data science can prevent tax evasion and predict incarceration rates.
- **E-commerce:** Data science can automate digital ad placement.
- **Gaming:** Data science can improve online gaming experiences.
- **Social media:** Data science can create algorithms to pinpoint compatible partners.
- **Fintech:** Data science can help create credit reports and financial profiles, run accelerated underwriting and create predictive models based on historical payroll data.

Q5. Why Python so popular? Future of Python.

Python has been an immensely popular programming language for a long time. This is mainly due to its easy-to-learn syntax and versatility. You can use Python in almost any area of software development. For example, you can work with:

- Web development
- Data science
- Game development

Python is buzzing everywhere. From the career perspective also the future looks very promising for Python. Over 40,000 jobs are completely reserved for experienced Python software developers worldwide. It still has some things to work on in the future. They include

- speed
- compatibility with native environments
- runtime errors
- browser integrations

These things can cause problems while writing code in Python. But the benefits more than outweigh the weaknesses and make Python one of the most popular programming languages.

Adding to this, Python is also used in data analytics, Big Data, Machine Learning, Artificial intelligence, and other revolutionary technologies. It is expected to get more importance in the near future.

Q6. Why Python for data analytics?

Python is a multi-functional, maximally interpreted programming language with several advantages that are often used to streamline massive, and complex data sets.

Python has a number of distinguishing characteristics that make it the best option for data analysis.

- Easy to learn
- Flexibility
- Huge library's collection
- Graphics and visualization
- Built-in data analytics tools

Q7. Importance of Python in machine learning.

The practical implementation of Python in machine learning projects and tasks has made the work easier for developers, data scientists, and machine learning engineers. Python can be easily used to analyze and compose available data, which also makes it one of the most popular languages in data science.

Q8. Importance of Python in deep learning & NLP.

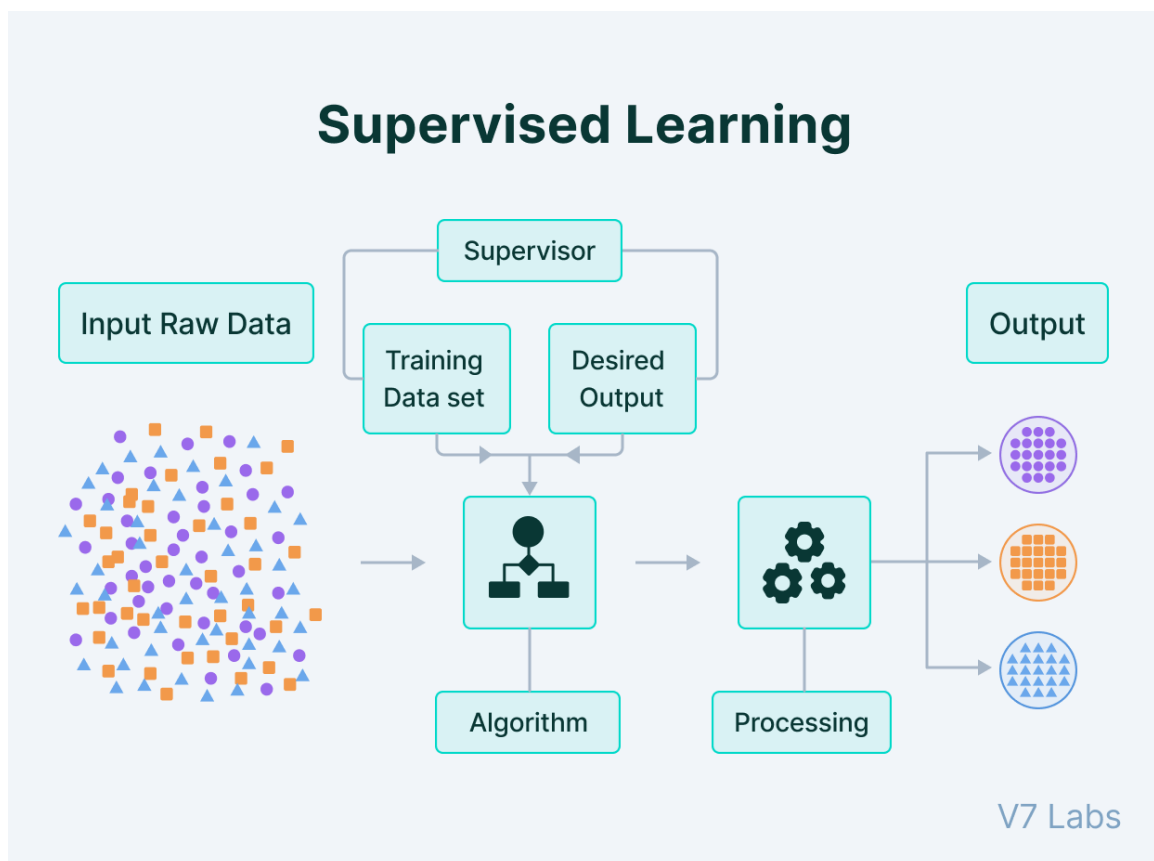
Deep learning is a subfield of machine learning involving artificial neural networks, which are algorithms inspired by the structure of the human brain. Deep learning has many applications and is used in many of today's AI technologies, such as self-driving cars, news aggregation tools, natural language processing (NLP), virtual assistants, visual recognition, and much more.

In recent years, Python has proven to be an incredible tool for deep learning. Because the code is concise and readable, it makes it a perfect match for deep learning applications. Its simple syntax also enables applications to be developed faster when compared to other programming languages. Another major reason for using Python for deep learning is that the language can be integrated with other systems coded in different programming languages. This makes it easier to blend it with AI projects written in other languages.

Python offers a versatile collection of NLP tools and libraries that enable developers to handle different NLP tasks, including sentiment analysis, POS tagging, document classification, topic modeling, word vectors, and more.

Q9. Supervised Learning with examples.

Supervised learning, as the name indicates, has the presence of a supervisor as a teacher. Basically, supervised learning is when we teach or train the machine using data that is well labelled. Which means some data is already tagged with the correct answer. After that, the machine is provided with a new set of examples(data) so that the supervised learning algorithm analyses the training data(set of training examples) and produces a correct outcome from labelled data.



Supervised Machine Learning Applications

Now, let's have a look at some of the popular applications of Supervised Learning:

- Predictive analytics (house prices, stock exchange prices, etc.)
- Text Recognition
- Spam detection
- Customer sentiment analysis
- Object detection (e.g. face detection)

Q10. Unsupervised learning with examples.

Unsupervised Learning is a type of machine learning in which the algorithms are provided with data that does not contain any labels or explicit instructions on what to do with it. The goal is for the learning algorithm to find structure in the input data on its own.

To put it simply—Unsupervised Learning is a kind of self-learning where the algorithm can find previously hidden patterns in the unlabeled datasets and give the required output without any interference.

Identifying these hidden patterns helps in clustering, association, and detection of anomalies and errors in data.

Applications

Supervised Learning models are ideal for classification and regression in labeled datasets. Spam detection, image classification, weather forecasting, price prediction is among their most common applications.

Unsupervised Learning fits perfectly for clustering and association of data points, used for anomaly detection, customer behavior prediction, recommendation engines, noise removal from the dataset, etc.

Machine Learning

Unsupervised Learning

Dimensionality Reduction

- Feature Elicitation
- Meaningful Compression
- Structure Discovery
- Big data visualization

Clustering

- Recommender Systems
- Targeted Marketing
- Customer Segmentation

Supervised Learning

Classification

- Identity Fraud Detection
- Image Classification
- Customer Retention
- Diagnostics

Regression

- Population Growth Prediction
- Estimating life expectancy
- Market Forecasting
- Weather Forecasting
- Advertising Popularity Prediction

Reinforcement Learning

- Real-time decisions
- Game AI
- Robot Navigation
- Learning Tasks
- Skill Acquisition

Q11. Role and responsibilities of a Database Engineer.

Assist in design and development of database systems.

Optimize database systems for performance and reliability.

Perform database maintenance and troubleshooting activities.

Test database systems and perform bug fixes.

Provide database solutions based on technical documents and business requirements.

Develop database functions, scripts, stored procedures and triggers to support application development.

Provide technical assistance to resolve all database issues related to performance, capacity and access.

Ensure data integrity and quality in database systems.

Maintain standard policies for database development activities.

Identify and rectify database errors in a timely manner.

Create physical and logical database models as per the business requirements.

Manage and monitor performance, capacity and security of database systems.

Prepare documentations regarding database design, configuration and change management tasks.

Mentor database administrators to manage the company databases effectively.

Perform data back-up and archival on regular basis.

Q12. Role and responsibilities of a Data Analyst.

A data analyst analyses data sets to find ways to solve problems relating to a business's customers. A data analyst also communicates this information to management and other stakeholders. The employment of these individuals encompasses many different industries such as business, finance, criminal justice, science, medicine, and government.

The role of a data analyst can be defined as someone who has the knowledge and skills to turn raw data into information and insight, which can be used to make business decisions.

Q13. Role and responsibilities of a Data Engineer.

Data engineers are mainly tasked with transforming data into a format that can be easily analyzed. They do this by developing, maintaining, and testing infrastructures for data generation. Data engineers work closely with data scientists and are largely in charge of architecting solutions for data scientists that enable them to do their jobs.

Q14. Role and responsibilities of a Data Scientist.

Data scientists work closely with business stakeholders to understand their goals and determine how data can be used to achieve those goals. They design data modeling processes, create algorithms and predictive models to extract the data the business needs, and help analyze the data and share insights with peers. While each project is different, the process for gathering and analyzing data generally follows the below path:

1. Ask the right questions to begin the discovery process
2. Acquire data
3. Process and clean the data
4. Integrate and store data
5. Initial data investigation and exploratory data analysis
6. Choose one or more potential models and algorithms
7. Apply data science techniques, such as machine learning, statistical modeling, and artificial intelligence
8. Measure and improve results
9. Present final result to stakeholders
10. Make adjustments based on feedback
11. Repeat the process to solve a new problem

Q15. Role and responsibilities of a Machine Learning Engineer.

A machine learning engineer (ML engineer) is a person in IT who focuses on researching, building and designing self-running artificial intelligence (AI) systems to automate predictive models. Machine learning engineers design and create the AI algorithms capable of learning and making predictions that define machine learning (ML).

An ML engineer's primary goals are the creation of machine learning models and retraining systems when needed. Responsibilities vary, depending on the organization, but some common responsibilities for this role include:

- Designing ML systems.
- Researching and implementing ML algorithms and tools.
- Selecting appropriate data sets.
- Picking appropriate data representation methods.
- Identifying differences in data distribution that affects model performance.
- Verifying data quality.
- Transforming and converting data science prototypes.
- Performing statistical analysis.
- Running machine learning tests.
- Using results to improve models.
- Training and retraining systems when needed.
- Extending machine learning libraries.
- Developing machine learning apps according to client requirements.