**Experiment No 1 Roll\_No :**

**Aim :**

**Theory :**

Machine learning is a branch of artificial intelligence that enables algorithms to uncover hidden patterns within datasets. It allows them to predict new, similar data without explicit programming for each task. Machine learning finds applications in diverse fields such as image and speech recognition, natural language processing, recommendation systems, fraud detection, portfolio optimization, and automating tasks.

* **Need for Machine Learning :**

Machine learning is important because it allows computers to learn from data and improve their performance on specific tasks without being explicitly programmed. This ability to learn from data and adapt to new situations makes machine learning particularly useful for tasks that involve large amounts of data, complex decision-making, and dynamic environments.

* [**Predictive modeling**](https://www.geeksforgeeks.org/what-is-predictive-modeling/)**:**Machine learning can be used to build predictive models that can help businesses make better decisions. For example, machine learning can be used to predict which customers are most likely to buy a particular product, or which patients are most likely to develop a certain disease.
* [**Natural language processing**](https://www.geeksforgeeks.org/natural-language-processing-overview/)**:**Machine learning is used to build systems that can understand and interpret human language. This is important for applications such as voice recognition, chatbots, and language translation.
* [**Computer vision**](https://www.geeksforgeeks.org/computer-vision-introduction/)**:**Machine learning is used to build systems that can recognize and interpret images and videos. This is important for applications such as self-driving cars, surveillance systems, and medical imaging.
* [**Fraud detection**](https://www.geeksforgeeks.org/detecting-frauds-with-ml-and-ai/)**:** Machine learning can be used to detect fraudulent behavior in financial transactions, online advertising, and other areas.
* [**Recommendation systems**](https://www.geeksforgeeks.org/what-are-recommender-systems/)**:** Machine learning can be used to build recommendation systems that suggest products, services, or content to users based on their past behavior and preferences.

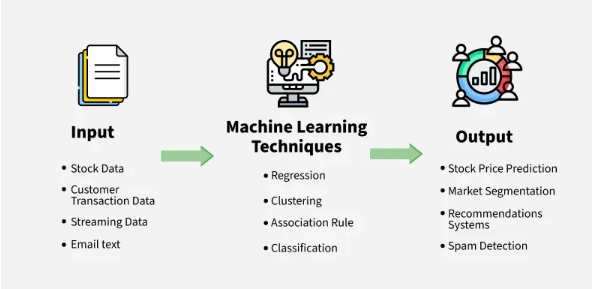


Fig 1.1

## How machine learning algorithms work?

### **1. Data Collection -**

First, relevant data is collected or curated. This data could include examples, features, or attributes that are important for the task at hand, such as images, text, numerical data, etc.

### **2. Data Preprocessing -**

Before feeding the data into the algorithm, it often needs to be preprocessed. This step may involve cleaning the data (handling missing values, outliers), transforming the data (normalization, scaling), and splitting it into training and test sets.

**3. Choosing a Model -**

Depending on the task (e.g., classification, regression, clustering), a suitable machine learning model is chosen. Examples include decision trees, neural networks, support vector machines, and more advanced models like deep learning architectures.

### **4. Training the Model -**

The selected model is trained using the training data. During training, the algorithm learns patterns and relationships in the data. This involves adjusting model parameters iteratively to minimize the difference between predicted outputs and actual outputs (labels or targets) in the training data.

### **5. Evaluating the Model -**

Once trained, the model is evaluated using the test data to assess its performance. Metrics such as accuracy, precision, recall, or mean squared error are used to evaluate how well the model generalizes to new, unseen data.

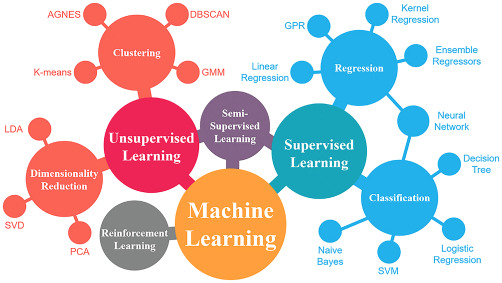
### **6. Fine-tuning -**

Models may be fine-tuned by adjusting hyperparameters (parameters that are not directly learned during training, like learning rate or number of hidden layers in a neural network) to improve performance.

### **7. Prediction or Inference -**

Finally, the trained model is used to make predictions or decisions on new data. This process involves applying the learned patterns to new inputs to generate outputs, such as class labels in classification tasks or numerical values in regression tasks.

* **Types of Machine Learning algorithms :**



## 1. Supervised Learning

### **A. Regression Algorithms** (Predict continuous values)

* **Linear Regression**  
  Simple model assuming a linear relationship between input and output.
* **Support Vector Regression (SVR)**  
  Uses margins (like SVM) for regression problems.
* **Decision Trees for Regression**
* **Random Forest Regressor**
* **Gradient Boosting Regressor**

### **B. Classification Algorithms** (Predict discrete labels)

* **K-Nearest Neighbors (KNN)**  
  Assigns class based on majority class of nearest neighbors.
* **Support Vector Machines (SVM)**  
  Finds the optimal hyperplane for separating classes.
* **Naive Bayes**  
  Based on Bayes theorem, assumes feature independence.
* **Decision Trees**
* **Random Forest**
* **Gradient Boosting Machines (XGBoost, LightGBM, CatBoost)**
* **Neural Networks (MLP, CNNs for image data)**

## 2. ****Unsupervised Learning****

### **A. Clustering Algorithms** (Group similar data points)

* **K-Means Clustering**
* **Hierarchical Clustering (Agglomerative, Divisive)**
* **DBSCAN (Density-Based Spatial Clustering)**
* **Gaussian Mixture Models (GMM)**

### **B. Dimensionality Reduction** (Reduce data features while retaining structure)

* **Principal Component Analysis (PCA)**
* **t-Distributed Stochastic Neighbor Embedding (t-SNE)**
* **UMAP (Uniform Manifold Approximation and Projection)**
* **Autoencoders (Neural Network-based)**

## 4. Reinforcement Learning (RL)

### **Key Algorithms:**

* **Q-Learning**
* **Deep Q Networks (DQN)**
* **SARSA**
* **Policy Gradient Methods (REINFORCE)**
* **Actor-Critic Methods**
* **Proximal Policy Optimization (PPO)**

## 5. Deep Learning Algorithms

### **Types:**

* **Convolutional Neural Networks (CNNs)** – image processing
* **Recurrent Neural Networks (RNNs)** – sequence modeling
* **Long Short-Term Memory (LSTM), GRU** – advanced RNNs
* **Transformers (BERT, GPT, etc.)** – NLP and beyond
* **Autoencoders** – feature learning & denoising
* **Generative Adversarial Networks (GANs)** – synthetic data generation

**Conclusion :**