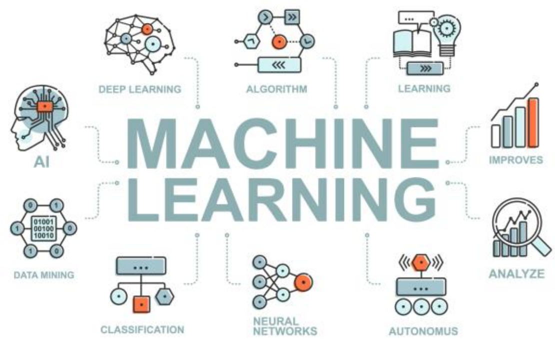
Introduction to Machine Learning

Concepts and Applications

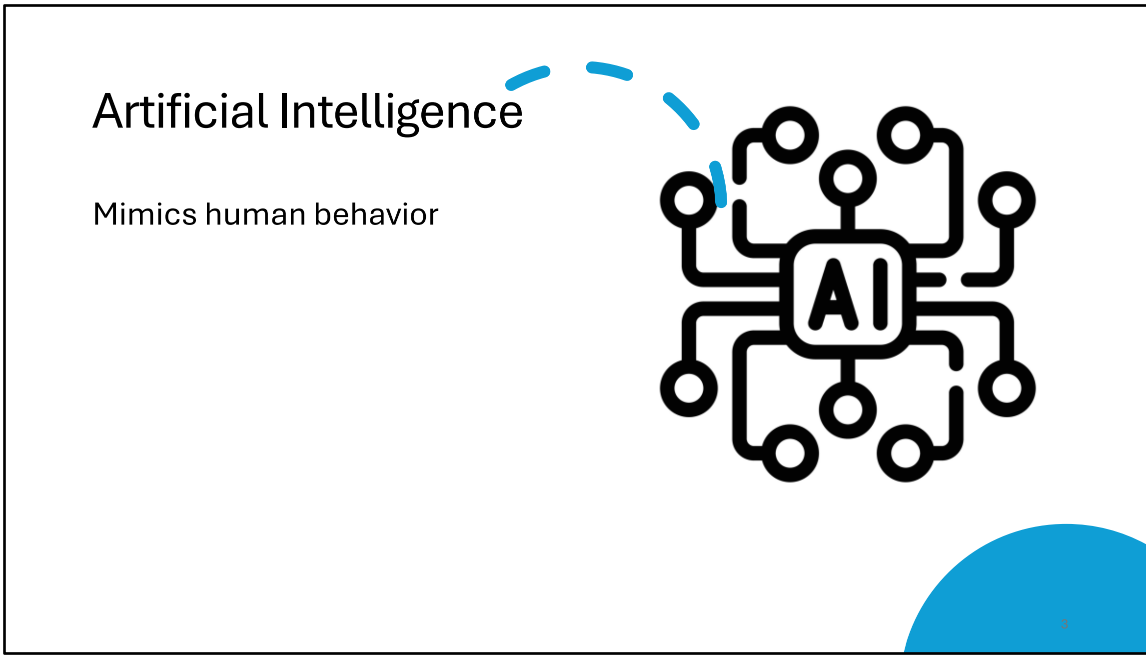


2

Session Objectives

- Fundamental concepts of AI and Machine Learning.
- Supervised, Unsupervised, and Reinforcement learning.
- Machine learning techniques and evaluation metrics.
- ML applications like Sentiment Analysis, Object Detection, Classification, Speech Recognition, NLP, and Chatbots.
- Practical considerations, ethical concerns.

These are the Topics we are going to cover in todays session.



A program/computer that mimics human behavior with human intervention.

It has the ability to make self decisions than just following the rule-based programming.

We see the Applications of AI and ML in everyday.

Where did you find the Application of ML/AI?

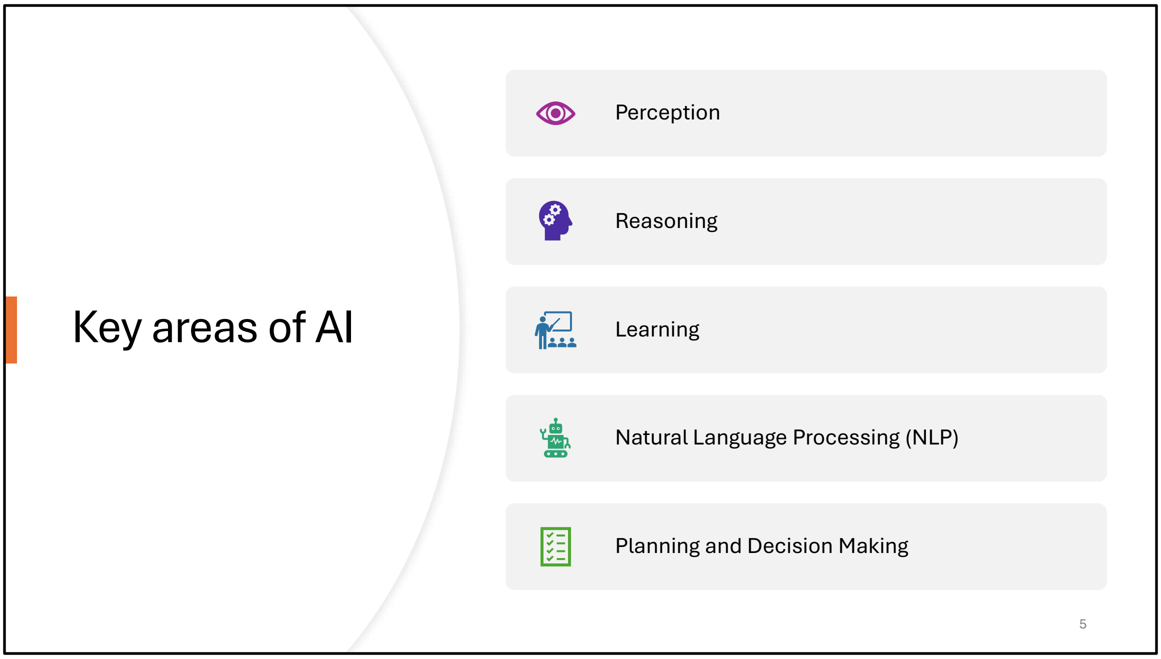
- Exam Proctoring
- Fraud transactions/Anomalies Detection in Industries IOT Sensors
- User Behavior/User Journey
- Loyal Customer(Contact Center Example)
- YouTube/Instagram Algorithm.
- Spamming

CLOUDEDGEIQ

4

Few of the places in

Exam Proctoring
Fraud transactions/Anomalies Detection in Industries IOT Sensors
User Behavior/User Journey
Loyal Customer(Contact Center Example)
Youtube/Instagram Algorithm.
Spaming



Perception:
•**Definition:** The ability of a system to interpret data from the world through sensory inputs.
•**Examples:**

- **Computer Vision:** Enabling machines to interpret and process visual information (e.g., facial recognition, object detection).
- **Speech Recognition:** Converting spoken language into text (e.g., virtual assistants like Siri, Alexa).

Reasoning:
•**Definition:** Drawing logical conclusions from data.
•**Examples:**

- **Expert Systems:** Using databases of expert knowledge to make decisions (e.g., medical diagnosis systems).
- **Game Playing:** AI strategies in games like chess or Go.

Learning:
•**Definition:** The ability of a system to improve performance based on previous experience.
•**Examples:**

- **Supervised Learning:** Training models with labeled data to make predictions.

- **Reinforcement Learning:** Learning through trial and error to maximize rewards (e.g., training robots, game AI).

Natural Language Processing (NLP):

•**Definition:** The ability of a machine to understand and generate human language.

•**Examples:**

- **Text Analysis:** Sentiment analysis, topic modeling.
- **Language Translation:** Translating text from one language to another (e.g., Google Translate).

How Translator works?

Rule based Translation: He's pulling my leg - वह मेरा पैर खींच रहा है

Neural Machine Translator: वह मजाक कर रहा है। (works by considering the broader context and linguistic nuances

(new-aun-cis)

Planning and Decision Making:

•**Definition:** The process of selecting actions to achieve specific goals.

•**Examples:**

- **Autonomous Vehicles:** Navigating routes, avoiding obstacles.
- **Robotics:** Task planning in dynamic environments.

Applications of AI

Healthcare

Finance

Retail

Transportation

Education

Entertainment

6

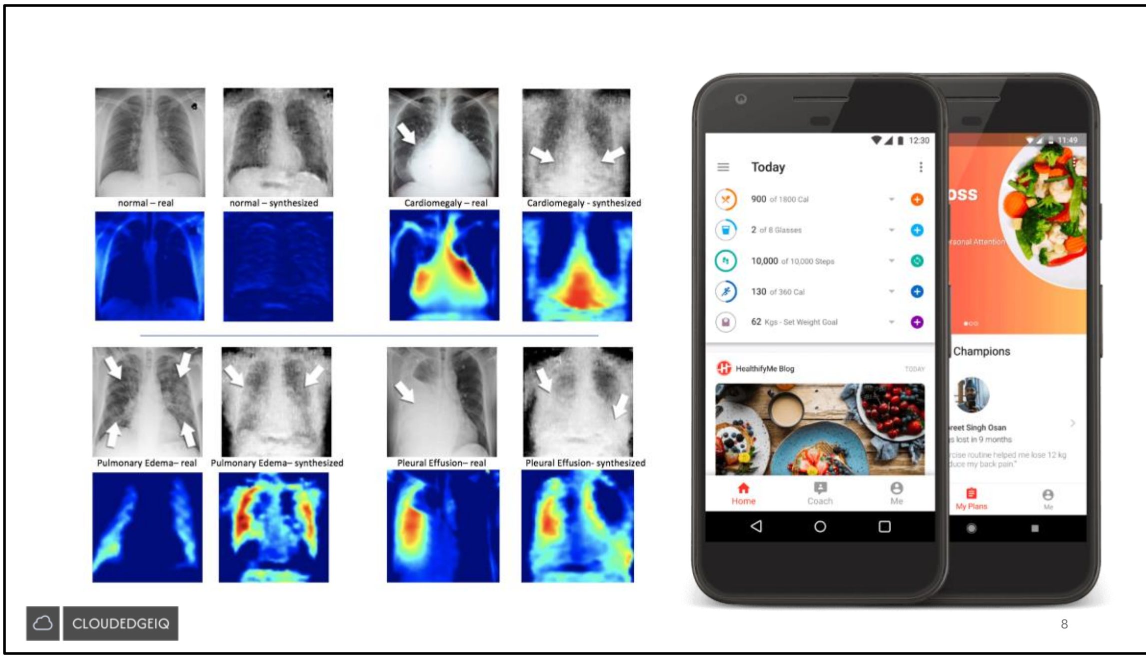
Next Slide

How in Healthcare?

- Predict disease outbreaks
- Diagnoses through image analysis (e.g., detecting tumors in MRI scans)
- Tailoring treatment plans based on patient data

7

Consider an X- Ray or UltraSonic Radiology Machine
Predicting Dengue Fever Outbreaks/
Heatwave Alert in mobilephone: using weather data.
AQI:
Personal Diet Applications; Like fitness,



How many of you have used these Health application

In Finance?

UNUSUAL PATTERNS IN TRANSACTION

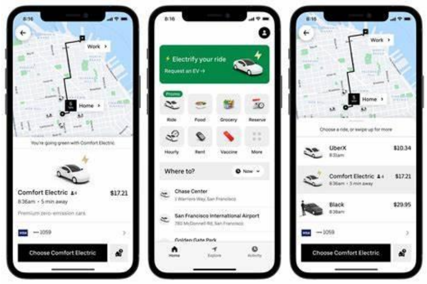
AI TO MAKE HIGH-SPEED TRADING DECISIONS

ANALYZING MARKET DATA TO ASSESS AND MITIGATE FINANCIAL RISKS

We use AI models to predict sales, so that the company can have the stock available and ready.

Adaptive Traffic Signal Control
uses
Reinforcement Learning: This technique allows the system to learn and improve its performance over time by receiving feedback from the environment

Deep Reinforcement Learning: An advanced form of RL that uses deep learning to handle more complex traffic scenarios.



CLOUDEDGEIQ

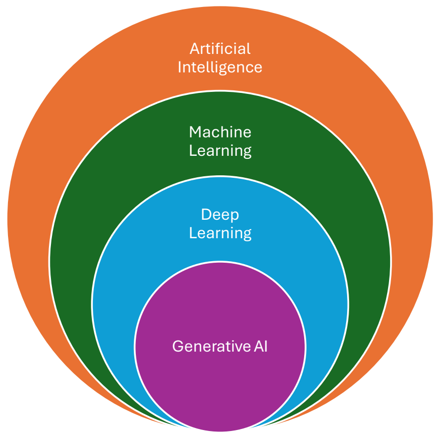
Maps use the AI to navigation and predict the traffic and duration of Journey
Uber use AI to predict the time to accept the trip by drivers.
Zomato use AI to help delivery boys to suggest locations to increase the probability of getting more orders.

Education and Entertainment?

- Personalizing educational content
- Assisting students with their queries
- Suggesting movies, shows, music
- Developing intelligent agents

13

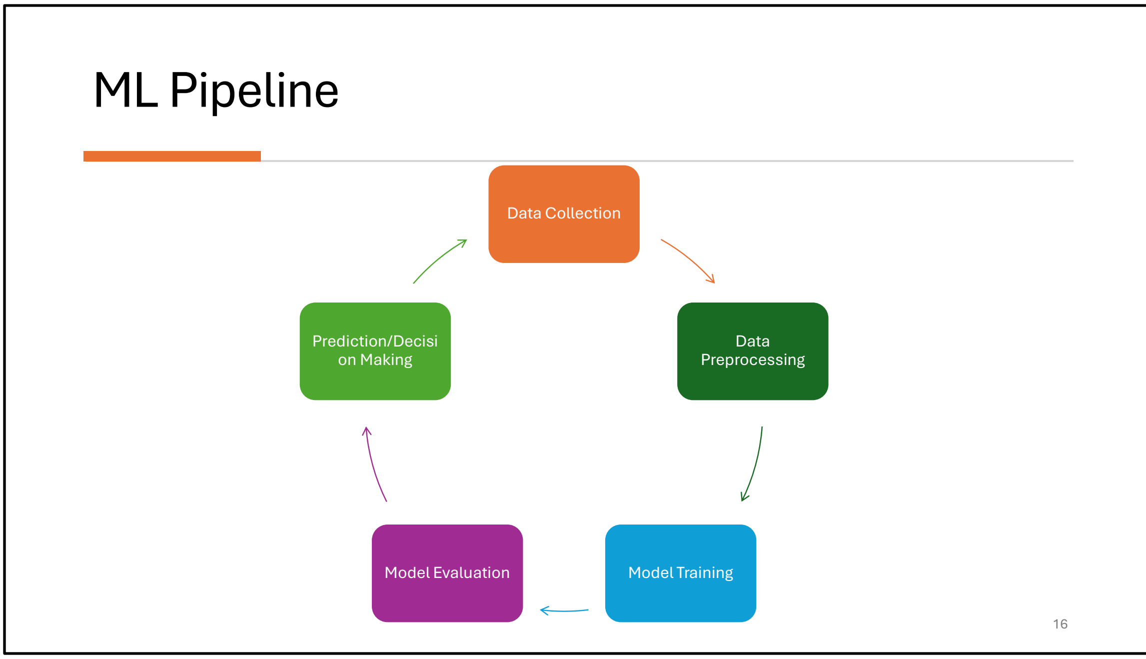
What is Machine Learning?



14

Machine Learning

- Subset of AI
- Developing algorithm from Data
- Make Predictions | Decisions

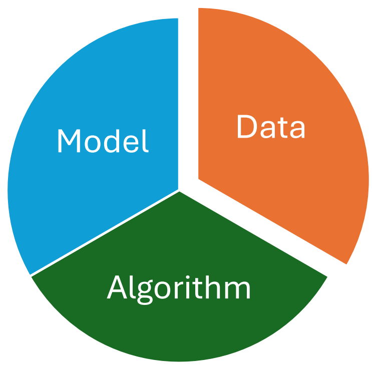


Data Collection: Gathering raw data from various sources.
Data Preprocessing: Cleaning and transforming the data to make it suitable for analysis.
Feature Extraction: Selecting and extracting relevant features from the preprocessed data.
Model Training: Training machine learning models on the processed data.
Model Evaluation: Assessing the performance of the trained models.
Model Deployment: Deploying the model into a production environment where it can make predictions on new data.

Example:
Data Collection: Gathering historical sales data, customer behavior data, and market trends.
Data Preprocessing: Cleaning the data, handling missing values, and normalizing features.
Feature Extraction: Selecting relevant features like purchase history, browsing behavior, and seasonal trends.

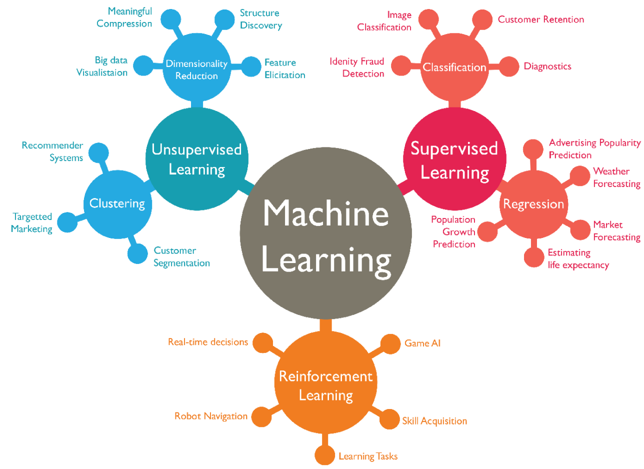
Model Training: Training a predictive model to forecast future sales based on the extracted features.
Model Evaluation: Evaluating the model's accuracy using test data.
Model Deployment: Deploying the model to predict sales in real-time and adjust inventory levels accordingly.

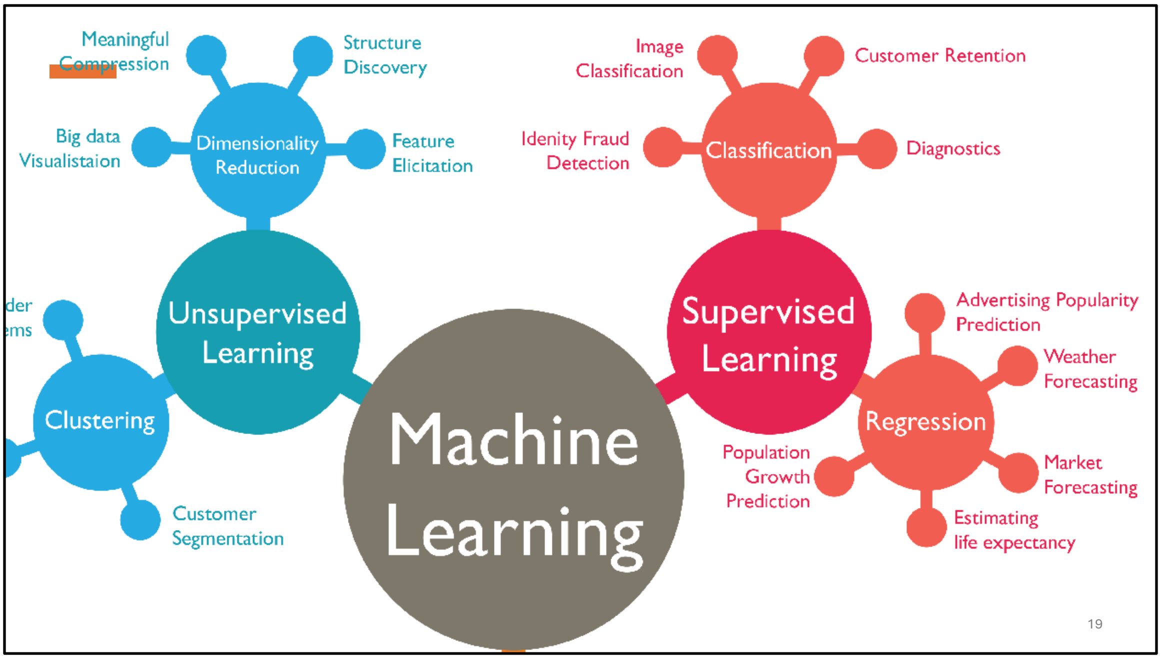
Components of ML



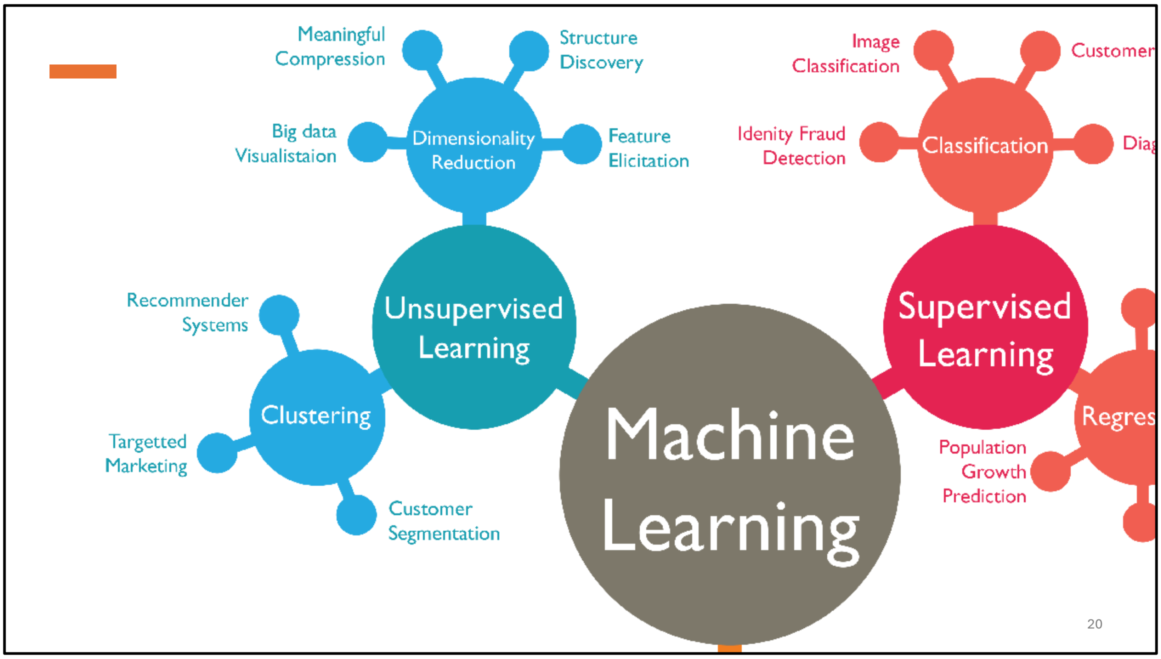
CLOUDEDGEIQ

Types of Machine Learning





Meaningful
Compression
Structure
Discovery
Big data
Visualistaion
Dimensionality
Reduction
Feature
Elicitation
Image
Classification
Customer Retention
Idenity Fraud
Detection
Classification
Diagnostics
Unsupervised
Learning
Supervised
Learning
Advertising Popularity
Prediction
Weather
Forecasting
Clustering
Regression
Machine
Learning
Population
Growth
Prediction
Market
Forecasting
Customer
Segmentation
Estimating
life expectancy
19

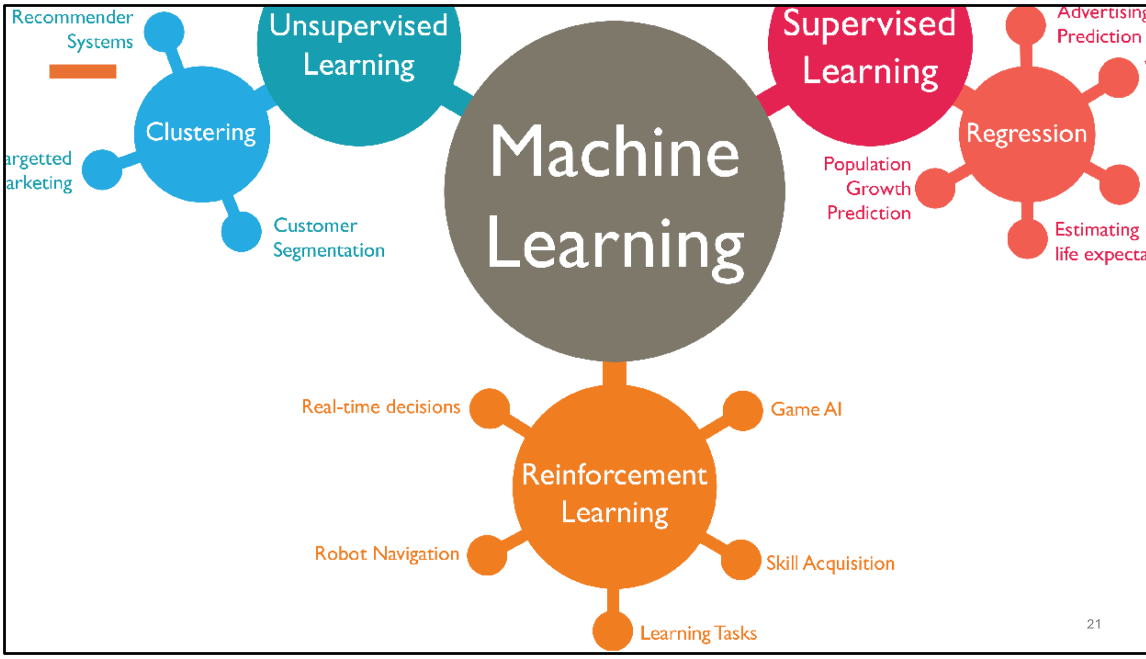


Customer Segmentation:
An e-commerce platform segments its customers based on their buying patterns. This helps in understanding the different types of customers they have and tailoring marketing strategies accordingly.

Identifies segments or groups within the customer base, such as "frequent buyers," "bargain hunters," or "fashion enthusiasts."

Target Marketing:
The same e-commerce platform uses the segments identified (e.g., "fashion enthusiasts") to send personalized emails showcasing new fashion arrivals, increasing the likelihood of conversions

Increases the effectiveness of marketing efforts by making them more relevant to the target audience.
Enhances customer engagement and boosts sales

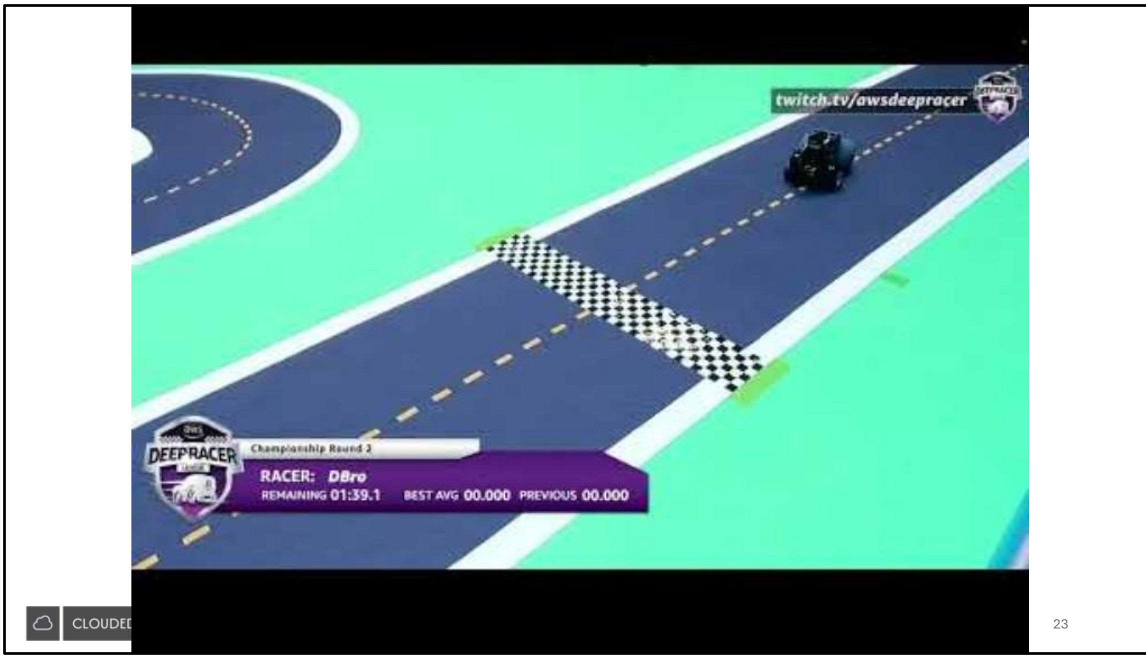

Machine Learning
Unsupervised Learning
Clustering
Recommender Systems
argetted arketing
Customer Segmentation
Supervised Learning
Regression
Advertising Prediction
Population Growth Prediction
Estimating life expectan
Reinforcement Learning
Real-time decisions
Game AI
Robot Navigation
Skill Acquisition
Learning Tasks
21

Reinforcement Learning

CLOUDEDGEIQ

22

AWS DeepRace,



Q-Learning: A value-based method where the agent learns the value of actions in each state.
Policy Gradient: A policy-based method where the agent directly optimizes the policy.
Deep Q-Networks (DQN): Combines Q-learning with deep neural networks to handle large state spaces.

Components of RL
Agent: The learner or decision-maker (e.g., a cleaning robot).
Environment: The external system the agent interacts with (e.g., the space the robot is cleaning).
State: A representation of the current situation or condition of the environment (e.g., the robot's position and the areas that need cleaning).
Action: The set of all possible moves the agent can make (e.g., moving forward, turning left, starting to clean).
Reward: A feedback signal indicating how good or bad an action is in a given state (e.g., a positive reward for cleaning an area, a negative reward for bumping

into obstacles).

The RL Process

1.Initialization:

1. The agent starts with no knowledge about the environment and initializes its policy (a strategy for choosing actions) and value function (a measure of how good a state or action is).

2.Interaction:

1. The agent observes the current state of the environment.
2. Based on its policy, the agent takes an action.
3. The environment responds to the action, transitioning to a new state and providing a reward.

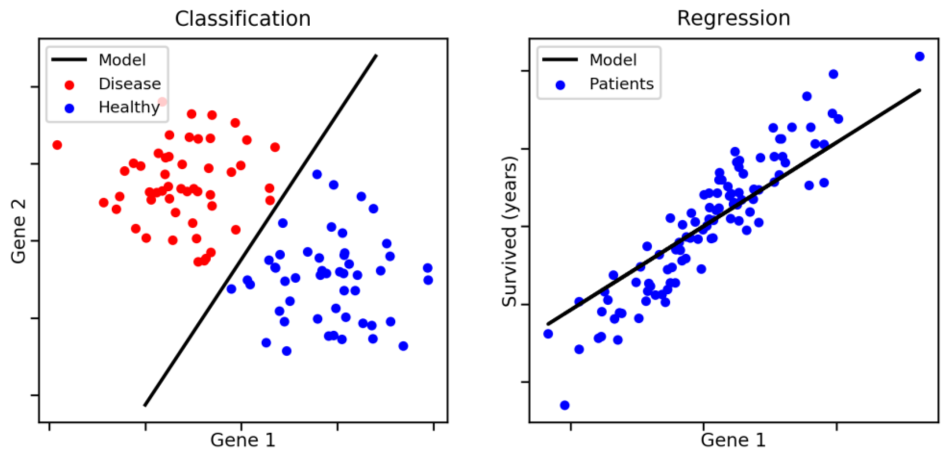
3.Learning:

1. The agent updates its policy and value function based on the reward received and the new state.
2. The goal is to improve the policy over time, choosing actions that maximize the cumulative reward.

4.Iteration:

1. This cycle of observation, action, reward, and learning continues until the agent's policy converges to an optimal strategy.

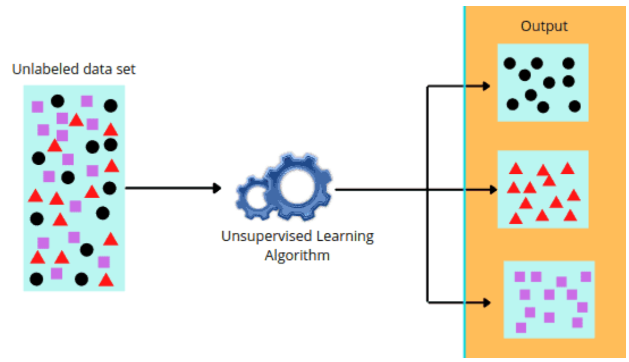
Supervised Learning



CLOUDEDGEIQ

Unsupervised Learning

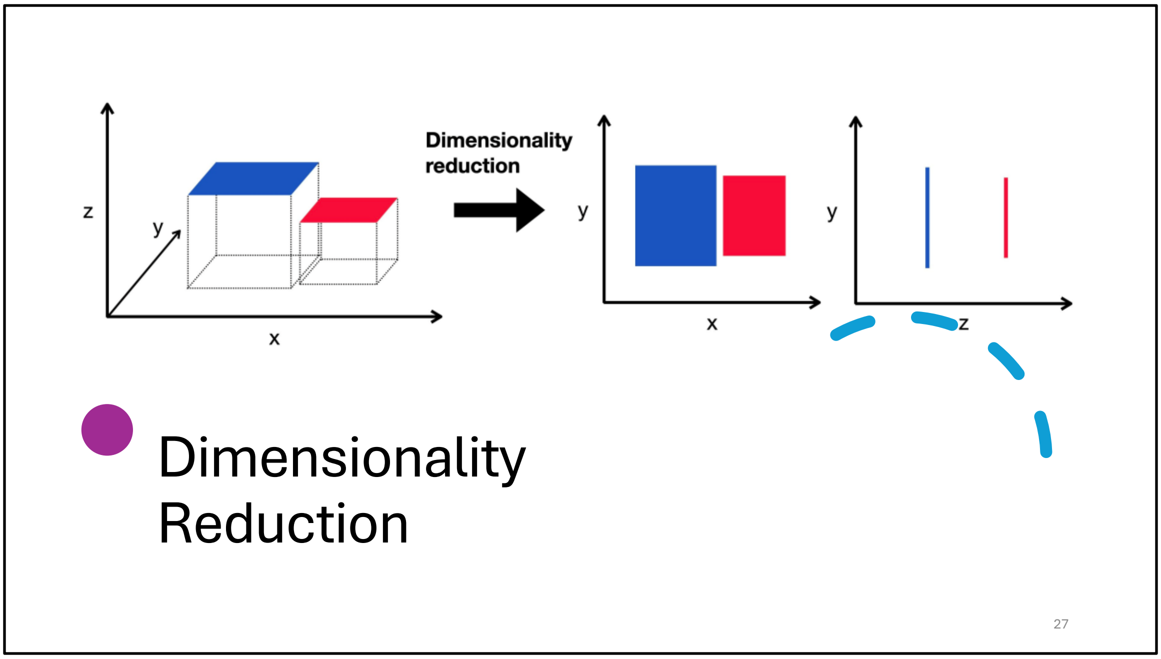
- Clusters
  - Grouping similar data points together
  - K-Mean Clustering
- Dimensionality Reduction:
  - Reducing the number of features while preserving important information



25

k-means be like:

Will learn more on Hands-on



Context: Many smartphones use facial recognition to unlock the device, authenticate transactions, or log into apps.

How It Works:

1.Data Collection: The smartphone's camera captures images of your face.

2.Feature Extraction: The system identifies key features of your face, such as the distance between your eyes, the shape of your nose, and other unique characteristics.

3.Dimensionality Reduction:

1. **Technique**: Methods like Principal Component Analysis (PCA) or Linear Discriminant Analysis (LDA) are used to reduce the high-dimensional feature space to a lower-dimensional one while preserving the most important information.
2. **Result**: This step simplifies the dataset, making it easier to process and match against stored facial templates.

4.Matching: The reduced-dimensional representation of your face is compared with stored templates to verify your identity.

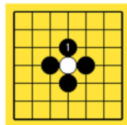
Reinforcement Learning

Learning through trial and error to maximize rewards in each environment

28

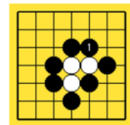
Developed by DeepMind,

Capturing Stones

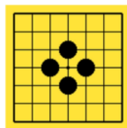


Dia 1-1: *Black plays 1*

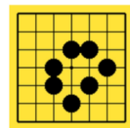
During a Go game one or more stones can be captured by completely surrounding them, i.e. filling all empty points around them. We show two examples: on the left a one stone capture, on the right a three stone capture. Once black has played 1 he removes the captured white stones from the board leading to the 2nd diagram positions.



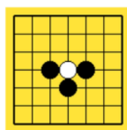
Dia 2-1: Black plays 1



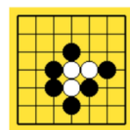
Dia 1-2: and takes one stone



Dia 2-2: and takes three stones!



Dia 1-3: Capturing one stone..



Dia 2-3: Capturing three stones...

Note 1: Needless to say that being captured usually is bad. There are numerous situations though where one would sacrifice one or more stones just to gain an advantage elsewhere.

CLOUDEDGEIQ

Self Decisions

When Not to Use RL

1. **Simple Problems**:
    1. For straightforward tasks where traditional supervised or unsupervised learning methods suffice.
2. **High Costs**:
    1. When the cost of exploration (trial and error) is too high, such as in critical medical applications without simulations.
3. **Limited Data**:
    1. When you don't have access to enough data to train the RL model effectively.

30

Autopilot and Tesla Cars use Reinforcement learning

Core Machine Learning Concepts

DATA PREPROCESSING:

ALGORITHMS AND MODELS:

EVALUATION METRICS:

MODEL OPTIMIZATION:

Data Preprocessing

1. Data Cleaning
2. Normalization
3. Feature Engineering
4. Data Augmentation

Quality and Consistency is crucial for model performance

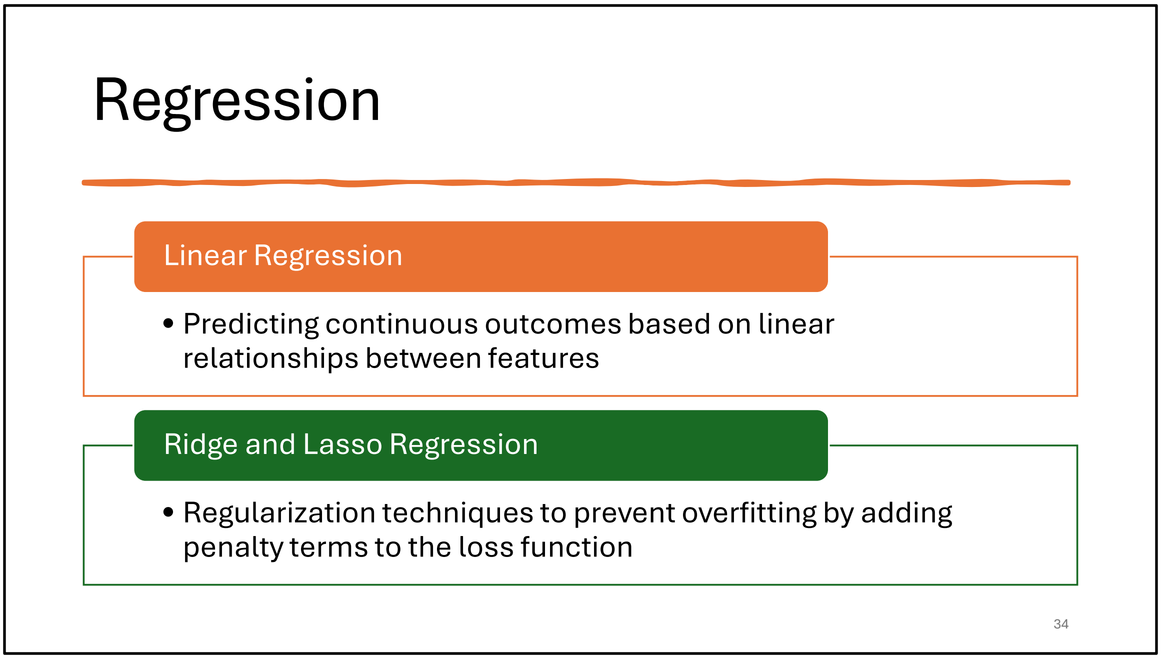
32

Data Cleaning: Handling missing values, outliers, and noise to improve data quality.
Normalization/Standardization: Scaling features to a similar range to ensure uniformity and improve model convergence.
Feature Engineering: Creating new features from raw data to enhance model performance.
Data Augmentation: Expanding the dataset by creating modified versions of existing data, especially useful for image and audio data.

Algorithms and Models

REGRESSION
ALGORITHMS

CLASSIFICATION
ALGORITHMS



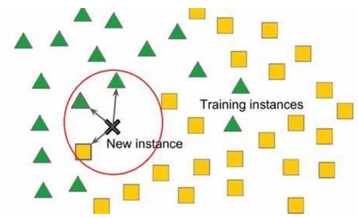
How It Works: Fits a linear equation (y = mx + c) to the observed data, minimizing the difference between the predicted and actual values.
Applications: Predicting house prices, sales forecasting, and other scenarios where the relationship between variables is linear.

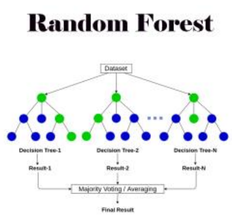
Ridge and Lasso Regression:
•**Purpose:** Regularization techniques to prevent overfitting by adding penalty terms to the loss function.
•**How They Work:**

- **Ridge Regression (L2 Regularization):**
  - Adds a penalty equivalent to the square of the magnitude of coefficients.
  - Helps prevent overfitting by shrinking the coefficients.
- **Lasso Regression (L1 Regularization):**
  - Adds a penalty equivalent to the absolute value of the magnitude of coefficients.
  - Can lead to sparse models, where some coefficients are exactly zero, effectively performing feature selection.

Applications: Useful in high-dimensional datasets where overfitting is a concern, such as genomics and financial modeling.

Classification





Logistic Regression

- For Binary Classification

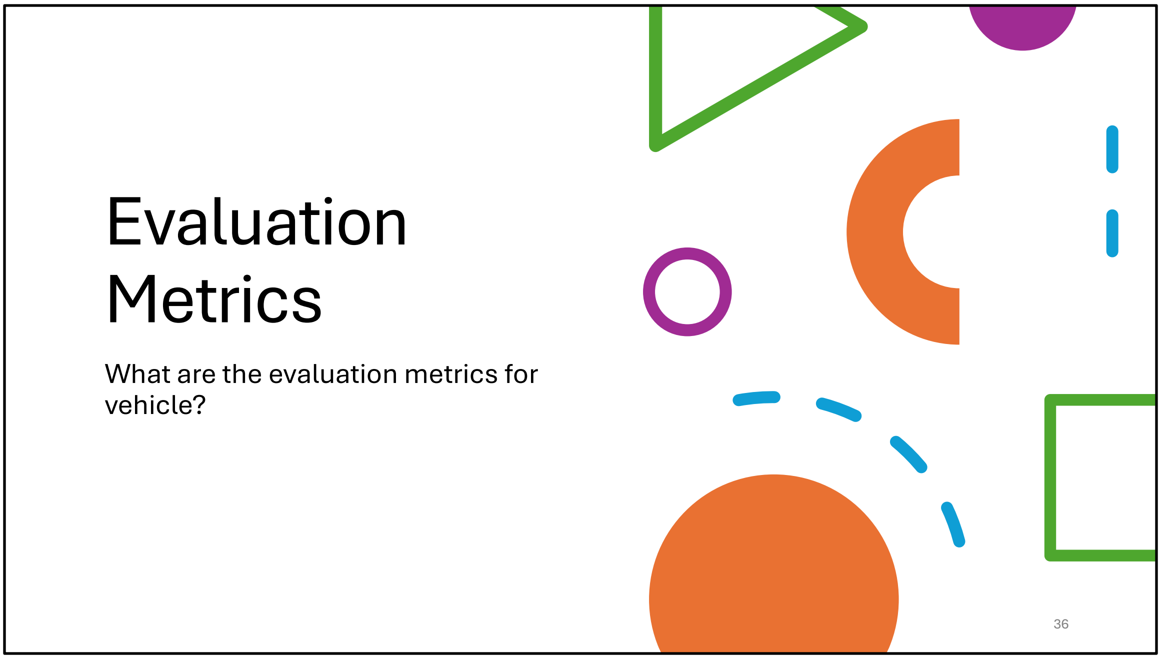
K-Nearest Neighbors (KNN)

- It classifies data points based on the labels of their nearest neighbors in the feature space

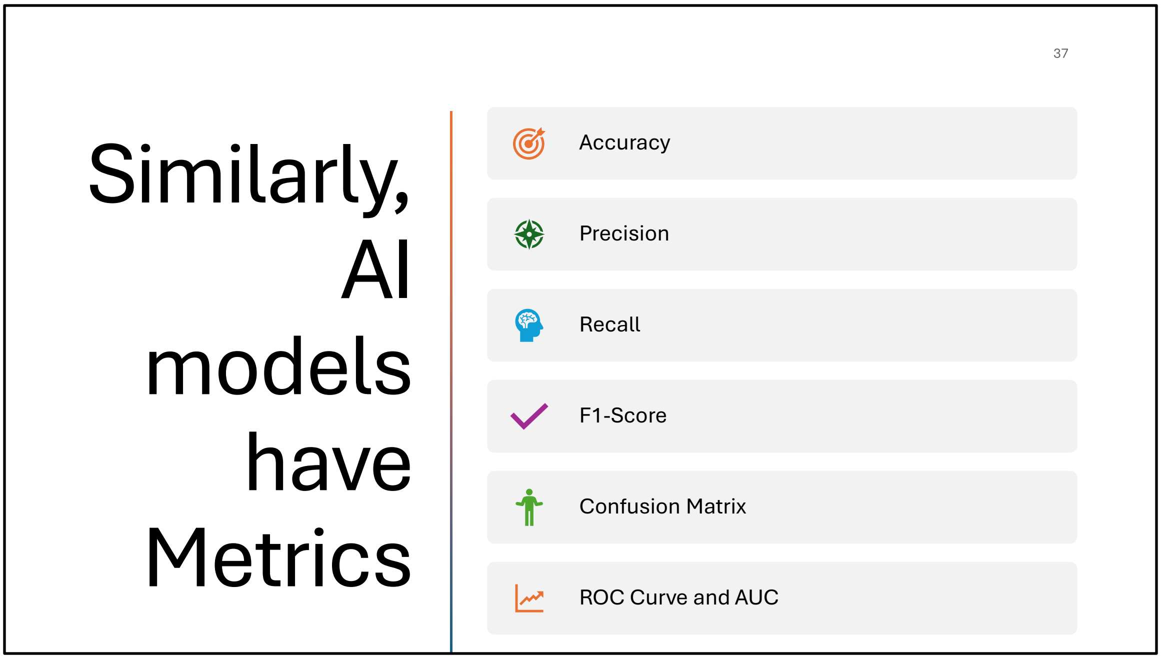
Random Forests

- This algorithm builds multiple decision trees and merges them to get a more accurate and stable prediction

CLOUDEDGEIQ



Mileage (Fuel Efficiency): Indicates how far a vehicle can travel on a given amount of fuel. It's crucial for understanding the cost-efficiency and environmental impact.
Speed: Measures how fast a vehicle can travel. It's important for performance and usability, especially in certain contexts like racing or commuting.
Drag (Aerodynamic Efficiency): Refers to the resistance a vehicle encounters as it moves through the air.
Acceleration: Measures how quickly a vehicle can increase its speed, typically from 0 to 60 mph (or km/h). This is important for performance and safety.
Braking Distance: Indicates the distance a vehicle takes to come to a complete stop from a certain speed. It's crucial for safety.
Horsepower and Torque: Metrics related to the engine's power and the vehicle's ability to perform tasks like towing.



Accuracy: Ratio of correctly predicted instances to total instances.
Precision: Ratio of correctly predicted positive instances to total predicted positive instances.
Recall: Ratio of correctly predicted positive instances to all actual positive instances.
F1-Score: Harmonic mean of precision and recall.
Confusion Matrix: Table comparing actual and predicted labels with true/false positives/negatives.
ROC Curve: Graph of true positive rate vs. false positive rate across thresholds.
AUC: Area under the ROC curve, measuring overall model performance.

Evaluation Metrics

Accuracy:

Correct predictions among total predictions

Precision:

True positive predictions among all positive predictions

	Predicted: Spam	Predicted: Not Spam
Actual: Spam	70	30
Actual: Not Spam	20	80

38

Confusion Matrix

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

Confusion Matrix

	Predicted 0	Predicted 1
Actual 0	70	30
Actual 1	20	80

Evaluation Metrics

Accuracy

$$\frac{TP + TN}{Total\ instances}$$

$$\frac{70+80}{70+30+20+80} = \frac{150}{200} = 0.75$$

Precission

$$\frac{TP}{TP + TN}$$

$$\frac{70}{70 + 20} = \frac{70}{90} = 0.78$$

Evaluation Metrics

Recall

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} = \frac{70}{70+30} = \frac{70}{100} = 0.7$$

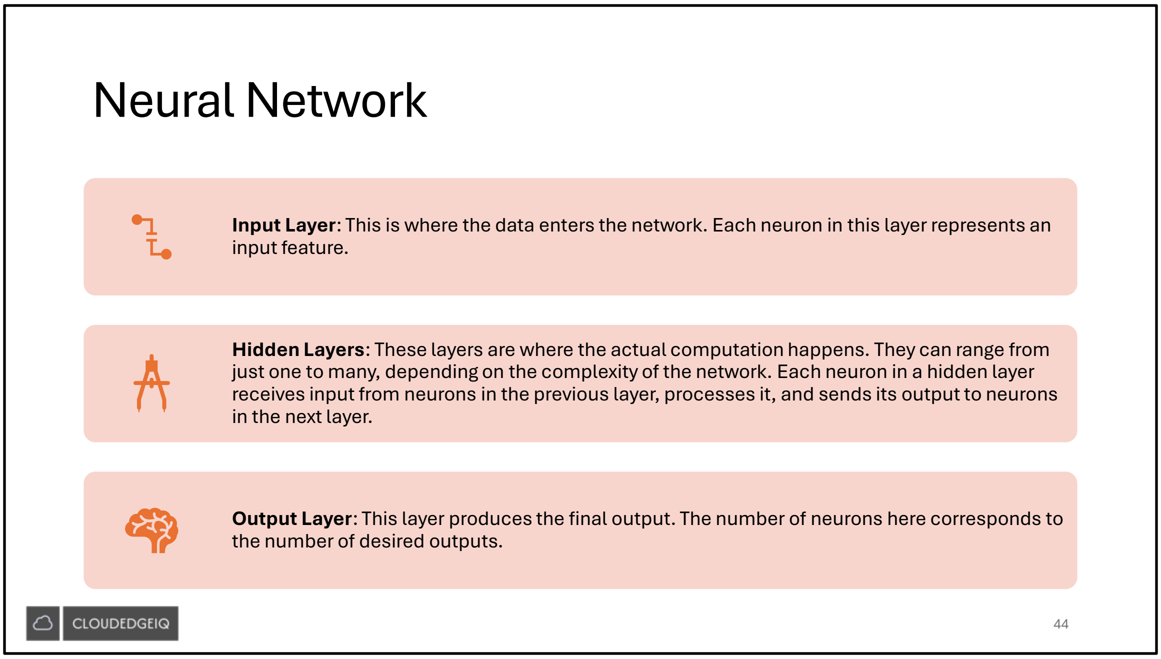
F1 Score

$$\text{F1-Score} = 2 \times \frac{\text{Precision}\times\text{Recall}}{\text{Precision+Recall}} = 2 \times \frac{0.78 \times 0.7}{0.78 + 0.7} \approx 0.74$$

70-80%: Adequate, may need improvement based on use case.
80-90%: Good, suitable for most applications.
90-95%: Very good, high-performing model.
95-98%: Excellent, indicates a highly accurate model.

Deep Learning Fundamentals

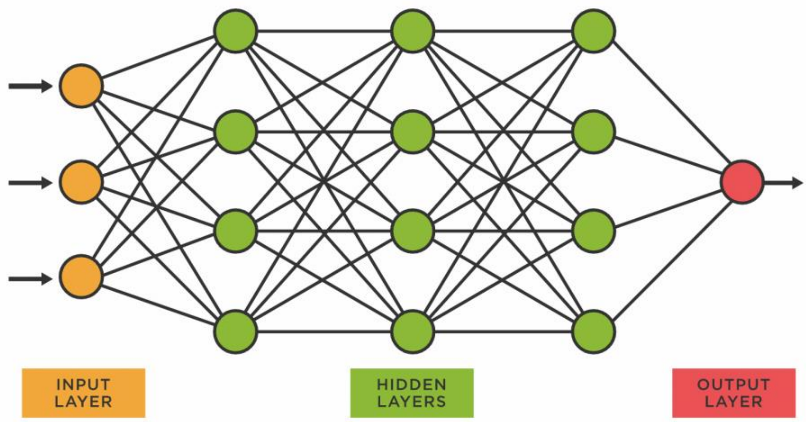
- NEURAL NETWORKS
- CONVOLUTIONAL NEURAL NETWORKS (CNNS)
- RECURRENT NEURAL NETWORKS (RNN'S)
- TRANSFORMERS



They consist of interconnected layers of nodes (or neurons) that work together to process and learn from data.

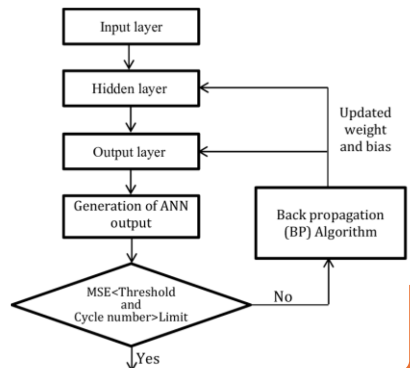
Weights and Biases: Each connection between neurons has an associated weight, and each neuron has a bias. These parameters are adjusted during training to optimize the network's performance.
Activation Functions: Functions applied to the output of each neuron to introduce non-linearity and enable the network to learn complex patterns (e.g., ReLU, Sigmoid, Tanh).

Neural Network Architecture



Steps in Neural Network

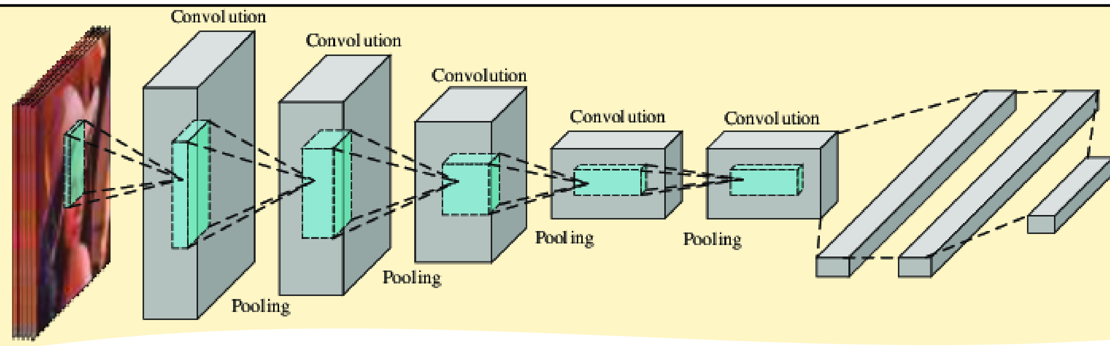
- Input Data
- Forward Pass
- Activation Functions
- Output
- Loss Calculation
- Backpropagation



46

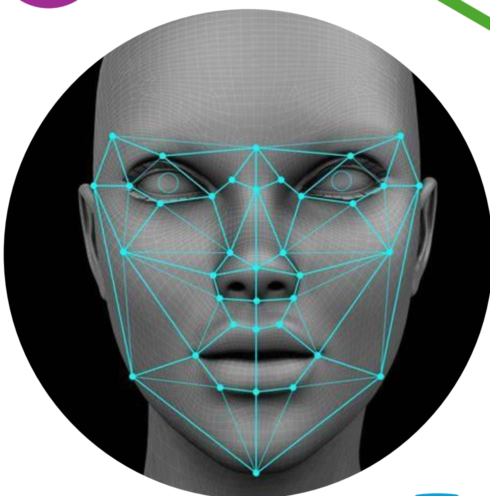
Activation Functions:
•Functions applied to the output of each neuron to introduce non-linearity. Common activation functions include ReLU (Rectified Linear Unit), Sigmoid, and Tanh.
•**ReLU**: Outputs the input directly if positive, otherwise, it outputs zero.
•**Sigmoid**: Compresses the input to the range (0,1).
•**Tanh**: Compresses the input to the range (-1,1).

Forward Propagation:
•The process of passing inputs through the network, layer by layer, to obtain the output.
Backpropagation:
•The process of adjusting weights and biases in the network based on the error (difference between predicted and actual output) to minimize the loss function



Then what are CNN?

CNNs are specialized types of neural networks designed to process and analyze visual data, such as images and videos



Components of CNN

Convolutional Layer

Pooling Layer

Activation Function

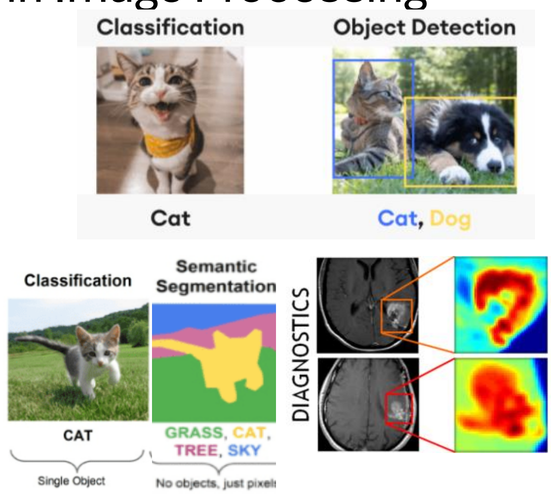
Fully Connected Layer

48

We uses this in our Faceial Recognization.
Snapchat.

How CNN helps in Image Processing

- Image Classification
- Object Detection
- Image Segmentation
- Face Recognition
- Medical Imaging

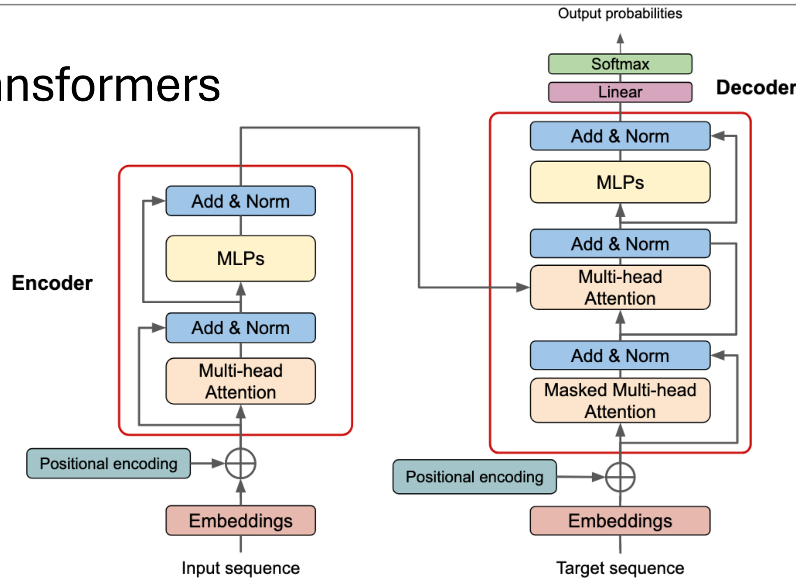


CLOUDEDGEIQ

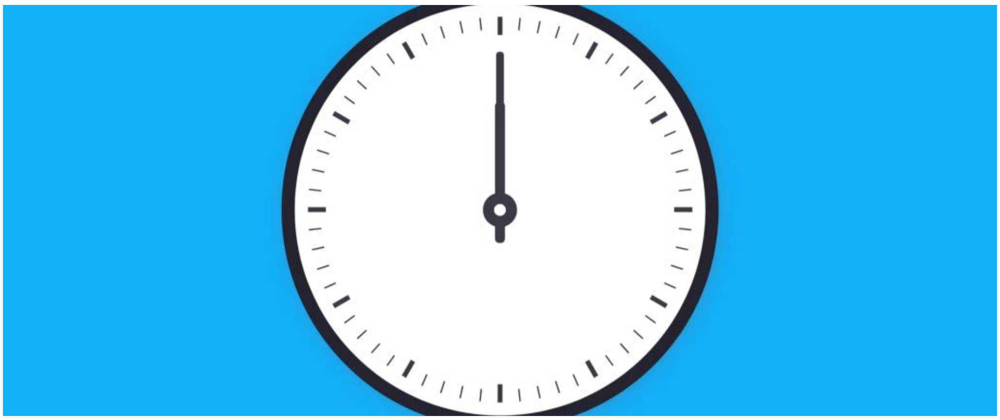
49

1. **Image Classification:**
    1. Assigning labels to images (e.g., cat, dog, car).
2. **Object Detection:**
    1. Identifying and localizing objects within an image.
3. **Image Segmentation:**
    1. Dividing an image into meaningful regions (e.g., separating foreground from background).
4. **Face Recognition:**
    1. Identifying and verifying individuals in images.
5. **Medical Imaging:**
    1. Analyzing X-rays, MRIs, and CT scans for diagnostic purposes.

Transformers



Its Handson Time



Hands-on

1. Sentiment Analysis
2. Placement Prediction Model(Classification Model)
3. Clustering
4. Vision

github link: https://github.com/demo-helloworld-1/MLplayground

CLOUDEDGEIQ

Data Privacy and Ethical Considerations

Data protection is critical to safeguard personal and sensitive information.

Ethical AI ensures responsible development and deployment of AI technologies.

Compliance with legal regulations (e.g., GDPR, CCPA) is mandatory.

Implementing privacy-preserving techniques like anonymization and encryption.

Transparent data usage policies build trust with users.

Bias and Fairness in ML

BIAS CAN ARISE FROM DATA COLLECTION, ALGORITHM DESIGN, AND HUMAN INTERVENTION.

REGULARLY AUDIT DATASETS FOR REPRESENTATIVENESS AND BALANCE.

IMPLEMENT FAIRNESS-AWARE ALGORITHMS TO MITIGATE BIAS.

USE TECHNIQUES LIKE RE-SAMPLING, RE-WEIGHTING, AND ADVERSARIAL DEBIASING.

ENGAGE DIVERSE TEAMS TO REVIEW AND EVALUATE MODEL OUTCOMES.

-Thankyou

55