

Machine Learning

* ML

→ Subset of AI

→ Training algo to automatically learn patterns

* Types of ML

→ supervised : labelled data

→ Unsupervised : Unlabelled data

→ Reinforced : Feedback learning

* Applications of ML:

→ Natural Language Processing

→ Fraud detection

→ Cyber security

→ Health diagnoses

→ Threat detection

→ Forecasting

* Types of problems in supervised learning:-

1) Regression ~~prob~~ problems

→ Predicting a continuous value

→ Goal → To learn mapping function

2) Classification problems

* Algorithms used:

Supervised

Linear regression

Logistic regression

Decision Trees

Neural networks

Unsupervised

K-means clustering

Association Rule mining

Principal Component Analysis

Hierarchical clustering

* Evaluation of performance :-

Supervised

accuracy

Precision

Recall

F₁ score

Area under ROC curve

Unsupervised

Using internal evaluation metrics:

→ Clustering → separation

Using external evaluation metrics

→ Purity → Entropy

* Deep learning:

→ subset of ML

→ Use artificial neural networks to extract features from data & to make predictions

* Deep learning models:

→ CNN: Convolutional Neural Network

→ RNN: Recurrent Neural Network

→ GAN: Generative Adversarial Network

* Applications of DL:-

→ self-driving car

→ NLP

→ speech recognition

→ medical image analysis

* Neural Network: ML model inspired by structure & function of human brain. Consists interconnected neurons to transmit info.

* Types of layers in neural network:

→ Input layer

→ Hidden layer

→ Output layer

* Backpropagation:

Algo used to train Neural Networks by updating weights & biases of neurons.

* Feature selection:

→ selecting most relevant features from dataset

→ Importance:

→ Reduces dimensionality

→ Prevents overfitting

→ Speeds up training time

* Feature engineering:

→ creating new features from existing ones.

→ Importance:

→ Captures more meaningful info

→ Improves predicting power

* Techniques for: ~~feature engineering~~

Feature Engineering

→ Scaling

→ Normalization

→ Binning

Feature Selection

→ Filter methods

→ Wrapper methods

→ Embedded methods

* Definitions:

- 1) Accuracy \rightarrow percentage of correctly predicted instances
- 2) Precision \rightarrow Proportion of true positives out of all positives
- 3) Recall \rightarrow Proportion of true positives out of all actual trues
- 4) F1 score \rightarrow Harmonic mean of precision & recall

~~5)~~

* Overfitting

model becomes overly complex & fits data too closely

Prevention: cross-validation
regularization
early stopping

Underfitting

model becomes too simple & doesn't capture patterns in data

Prevention: Increase model complexity
collect more data

* Reinforcement learning: Make decisions through feedback in form of rewards & punishments. [Goal: Maximize reward]

* Applications

- \rightarrow Game playing
- \rightarrow Robotics
- \rightarrow Optimizing problems
- \rightarrow Recommendation systems

Algo used:

- \rightarrow Q-learning
- \rightarrow SARSA
- \rightarrow Deep reinforcement learning

* Performance Evaluation

- \rightarrow Computing cumulative reward s.