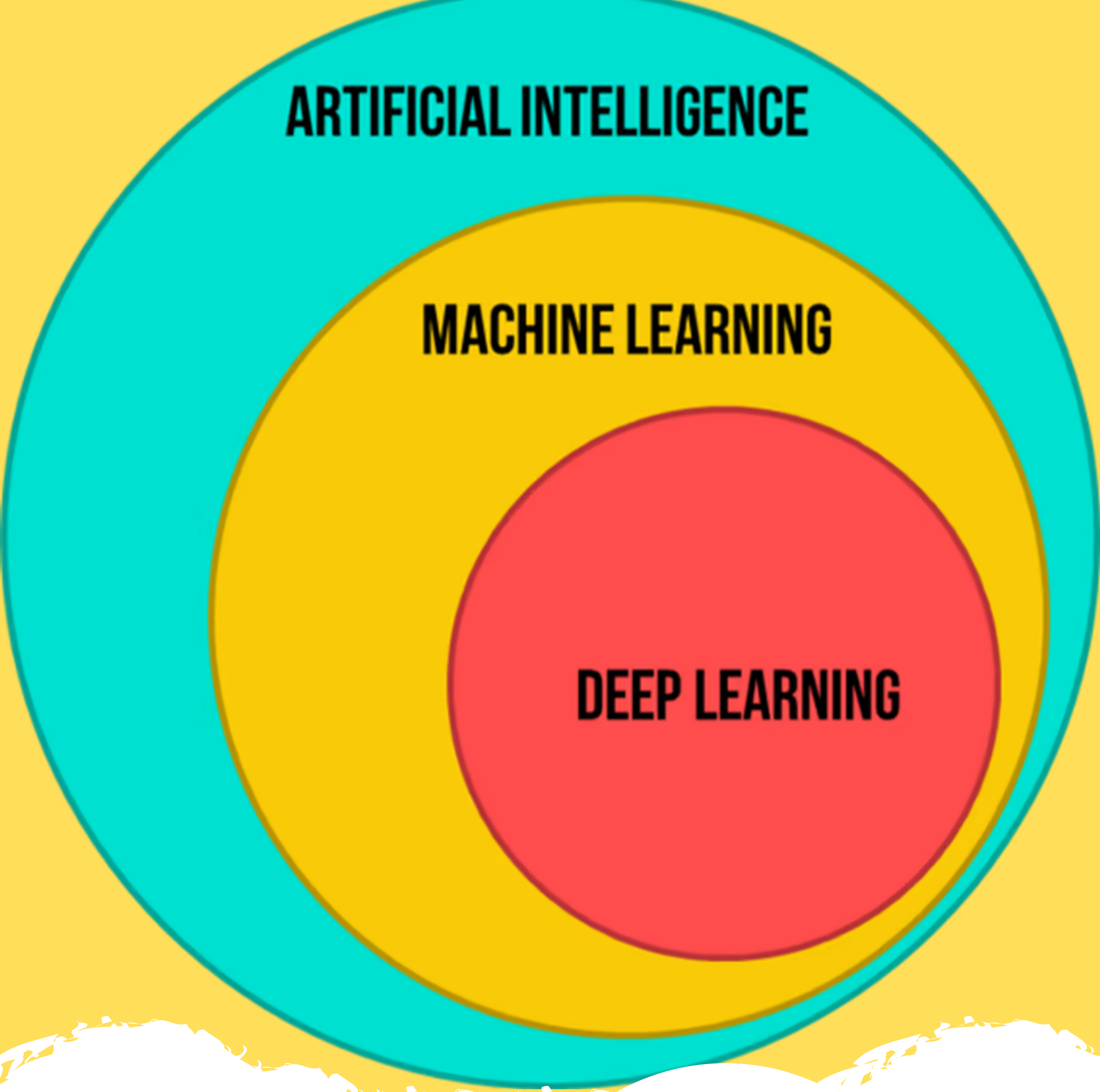
MACHINE LEARNING

Saptarshi Parui


ARTIFICIAL INTELLIGENCE
MACHINE LEARNING
DEEP LEARNING

DEFINATION:

Machine Learning is a type of computer technology that enables machines to learn and make predictions or decisions without being explicitly programmed. It's like teaching a computer to learn from examples and improve its performance over time.

MACHINE LEARNING

Supervised Learning

the computer is provided with labeled examples of input data and their corresponding desired outputs. It learns to recognize patterns and make predictions by mapping the input data to the correct output. For example, if given a dataset of labeled images, the algorithm can learn to classify new images based on the patterns it has observed.

Unsupervised Learning

the computer is given a dataset without pre-existing labels or categories. It aims to discover patterns, structures, or relationships in the data on its own. It doesn't have specific guidance on what to look for. For instance, unsupervised learning can be used to group similar customer behaviors in a dataset without any prior knowledge of customer segments.

Reinforcement Learning

the computer learns through interaction with an environment. It takes actions and receives feedback in the form of rewards or penalties. The objective is to learn the best actions to maximize cumulative rewards over time. For example, a reinforcement learning algorithm can learn to play games by exploring different actions and receiving rewards for winning or penalties for losing.

DATA

Labelled Data

Labelled data has specific tags or categories assigned to it. Each data point is marked with a known label that represents its desired output or category. For example, in a dataset of images, each image is tagged with labels like "cat," "dog," or "bird." Labelled data is used to train models in supervised learning to recognize patterns and make predictions based on the provided labels.

Unlabelled Data

Unlabelled data lacks specific tags or categories. It does not have any assigned labels representing its desired output. Unlabelled data contains raw or unlabeled information that requires further analysis. Unsupervised learning algorithms explore unlabelled data to discover hidden patterns, structures, or relationships without any predefined categories or labels. These algorithms group similar data points or identify clusters without prior knowledge of the data's inherent labels. Unlabelled data is also used in semi-supervised learning, where a small portion of the data is labelled, and the rest is unlabelled, to enhance the learning process.

MACHINE LEARNING ALGORITHMS

Supervised Learning Algorithms

- Linear Regression
- Logistic Regression
- Decision Trees
- Random Forest
- Support Vector Machines (SVM)
- Naive Bayes
- K-Nearest Neighbors (KNN)
- Gradient Boosting Methods (e.g., XGBoost, AdaBoost)
- Neural Networks (including Deep Learning)
- Ensemble Methods (e.g., Bagging, Stacking)

Unsupervised Learning Algorithms

- K-Means Clustering
- Hierarchical Clustering
- DBSCAN
- Gaussian Mixture Models (GMM)
- Principal Component Analysis (PCA)
- t-SNE (t-Distributed Stochastic Neighbor Embedding)
- Association Rule Learning (e.g., Apriori, FP-Growth)
- Self-Organizing Maps (SOM)
- Mean Shift Clustering
- Hierarchical Density-Based Spatial Clustering (HDBSCAN)

Reinforcement Learning Algorithms

- Q-Learning
- Deep Q-Networks (DQN)
- Policy Gradient Methods
- Proximal Policy Optimization (PPO)
- Actor-Critic Methods
- Monte Carlo Tree Search (MCTS)
- Temporal Difference Learning
- Deterministic Policy Gradient (DPG)
- Trust Region Policy Optimization (TRPO)
- Asynchronous Advantage Actor-Critic (A3C)

PROCESS OF MAKING A MACHINE LEARNING MODEL

LOAD DATASET

PREPROCESS DATA
(FEAUTURE ENGINEERING)

MODEL SELECTION

TRAIN YOUR MODEL

EVALUATION OF MODEL

DEPLOYMENT

EVALUATION TECHNIQUES FOR A MACHINE LEARNING MODEL

- Accuracy
- Confusion Matrix
- Precision
- Recall (Sensitivity or True Positive Rate)
- F1 Score
- ROC Curve (Receiver Operating Characteristic Curve)
- AUC-ROC (Area Under the ROC Curve)
- Mean Absolute Error (MAE)
- Mean Squared Error (MSE)
- Cross-Validation

PYTHON MODULES FOR MAKING A MACHINE LEARNING MODEL

- **NumPy: Mathematical operations and array manipulations for handling numerical data.**
- **Pandas: Data manipulation and analysis for structured data.**
- **Scikit-learn: Comprehensive machine learning library for classification, regression, clustering, and evaluation.**
- **TensorFlow: Deep learning library for building and training neural networks.**
- **Keras: High-level API for building deep learning models (runs on top of TensorFlow).**
- **PyTorch: Deep learning library with dynamic computation graphs.**
- **Matplotlib: Data visualization library for creating plots and charts.**
- **Seaborn: Statistical visualization library (built on top of Matplotlib).**
- **SciPy: Scientific computing tools for optimization, integration, linear algebra, and statistics.**
- **XGBoost: High-performance gradient boosting library for ensemble models.**

PLATFORMS FOR MAKING A MACHINE LEARNING MODEL

- JUPYTER NOTEBOOK
- GOOGLE COLLABORATORY
- KAGGLE
- VS CODE
- TENSORFLOW PLAYGROUND

THANK YOU!!!