Nikita Patra
21003015300067
AI ML - B
04 March 2024

MLT (Assignment 1)

4/3/24

Ques-1) Define machine Learning. Explain the types of machine Learning with examples.

Machine Learning is a subset of A.I.

Definition: According to Tom Mitchell a computer program is said to learn from experience E, with respect to some class of Task T and performance measure P, if its performance at tasks in T as measured by P, improves with experience E.

Machine learning can be broadly categorised into three main types: Supervised learning, unsupervised learning and reinforcement learning.

1. Supervised Learning: In supervised learning the algorithm learns from labelled data, which means it is provided with input-output pairs. The algorithm then learns a mapping function from the input to the output.

- Example:
  - Classification: Given a dataset of emails labeled as "spam" or "not spam", the algorithm learns to classify new e-mails as spam or not spam.
  - Regression: Predicting house prices based on features like size, location, no. of bedrooms etc.

2. Unsupervised Learning: In unsupervised learning, the algorithm learns pattern from unlabelled data. It does not have labelled outputs, so it finds the underlying structure or distribution in the data.

- Example:
  - Clustering: Grouping customers based on their purchasing behaviour without any prior knowledge of customer segments.
  - Dimensionality Reduction: Techniques like principle component analysis (PCA) or t distributed stochastic Neighbour embedding (t-SNE) for visualising high dimensional data.

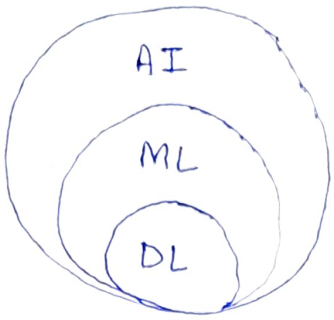
3. Reinforcement Learning : It is a type of machine learning where an agent learns to make decisions by taking actions in an environment to maximise some notion of cumulative reward. It learns through trial and error, receiving feedback from the environment in the form of rewards or penalties.

- Example
  - Game Playing : Training an AI to play video games such as chess or Go, where it learns through successive iterations to maximise its score.
  - Robotics : Teaching a robot to navigate a maze or perform tasks in a simulated or real world environment.

Ques-2) Discuss the history of Machine Learning.

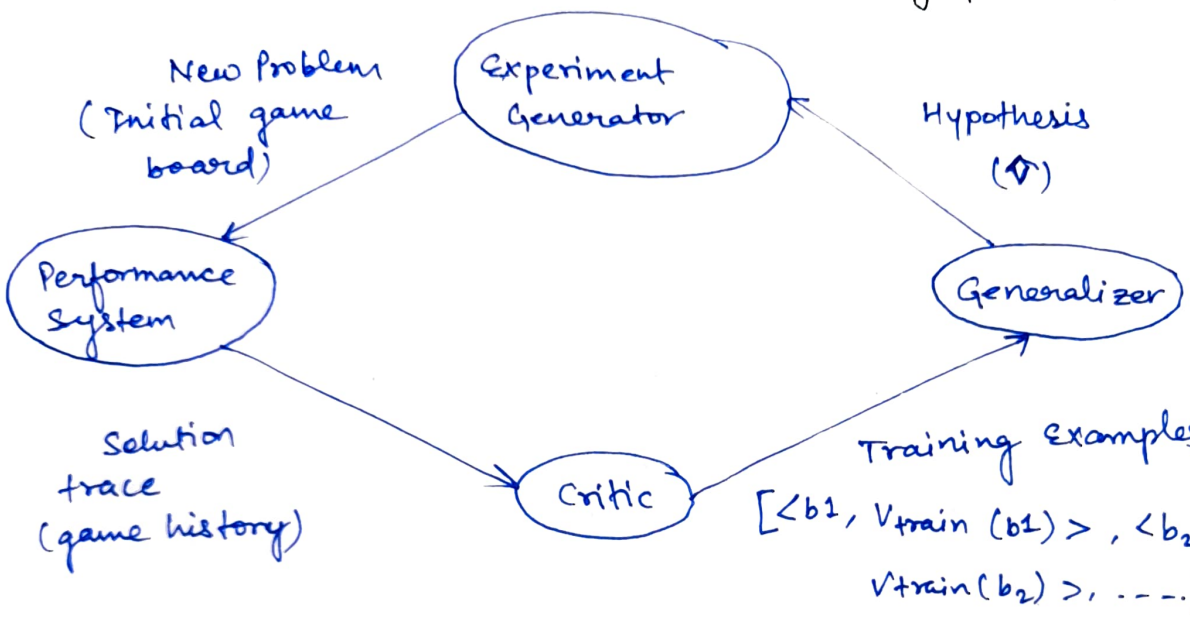
pre 1950s	Statistical method are discovered and redefined.
1950s	Pioneering machine learning research is conducted using simple algorithms.
1960s	Bayesian methods are introduced for probabilistic inference in machine learning.
1970s	AI winter caused by pessimism about machine learning effectiveness.
1980s	Rediscovery of back propagation causes a resurgence in machine learning research
1990s	Work on machine learning shifts from a knowledge-driven approach to a data-driven approach. Scientists begin creating programs for computer to analyze large amounts of data and draw conclusions - or "learn" - from the results. Support vector machines (SVMs) and recurrent neural networks (RNNs) become popular. The fields of computational complexity via neural networks and super-Turing computation started.
2000s	Support vector clustering and other kernel methods and unsupervised machine learning methods become widespread.
2010s	Deep learning becomes feasible, which leads to machine learning becoming integral to many widely used software services and applications.
2020s	Gen AI leads to revolutionary models, creating a proliferation of foundation models both proprietary and open source, notably enabling products such as chatgpt (text based) and stable Diffusion (Image based). Machine learning and AI enter the wider public consciousness.

Ques-3) What is a well posed learning problem. Explain with examples.



Venn Diagram of AI, ML & DL.

Ques-5) Give the final design of checker learning problem.



Performance measure: It's the module that must solve the given performance task. It takes new problem as input and produces a trace of its solution as output.

Critic: Critic takes as the input as the history or trace of the game and produces the output as a set of training examples of the target function Critic corresponds to the equation

$$V_{train}(b) \leftarrow \hat{V}(successor(b))$$

Generalizer: It takes as input the training examples and produces an output for hypotheses. It corresponds to LMS algorithm & output hypothesis is the function $\hat{V}$ described by the learning weights w_0, w_1, w_2, w_3

Tom Mitchell (1998) well-posed Learning problem: A co
program is said to learn from experience E with respe
to some task T and some performance measure P, if
its performance on T as measured by P, improves by
experience E.

for example:
Suppose your e-mail program watches you do or do not mark as spam and based on that learns how to better filter spam.

T: Classifying emails as spam or not spam
E: Watching you label emails as spam or not spam
P: The number (or fraction) of emails correctly classified as spam or not spam.

Ques-4) Differentiate Artificial Intelligence, Machine learning and Deep Learning.

Artificial Intelligence	Machine Learning	Deep Learning
(1) It is the study/ process which enables machine to mimic human behaviour through particular algorithm.	ML stands for machine learning and is the study that uses statistical methods enabling machines to improve with experience.	It is the study that makes use of Neural Networks (similar to neurons present in human brain) to imitate functionality just like a human brain.
(2) AI is the broader family consisting of ML and DL as its component.	ML is the subset of AI.	DL is the subset of ML.
(3) Computer algorithm which exhibits intelligence through decision making.	ML is an AI algorithm which allow system to learn from data.	DL is an ML algorithm that uses deep (more than one layer) neural networks to analyze data & provide output accordingly.

Experiment Generator: Takes as input the current hypothesis & output a new problem for the performance system to explore.

Ques-6) Discuss the various issues of machine learning techniques.

1. Insufficient training data (underfitting): Simple task requires thousands of sample of data whereas complex task like speech or Image processing requires millions of samples of data.

2. Poor quality of data:
Data quality is important for ideal Algorithm. Noisy and incorrect data is responsible innacurate predictions.

3. Algorithm selection:
For a particular task algorithm selection is difficult. It can be improved by comparing the errors derived from different algorithms.

4. Overfitting:
Overfitting means machine will have more samples for an object to classify and it includes noisy data. In this case the model will become complex. The main reason behind overfitting is non-linear methods used in Machine Learning Algorithms.

5. Monitoring & Maintenance:
Regular monitoring & maintenance is compulsory as different results for different actions require data change; hence editing of codes as well as ... for monitoring is necessary.

Data Bias:

Data biasing is a big challenge in machine learning. These error exists when certain elements of data set are heavily or weighted or need more importance from others. Biased data leads to inaccurate results.