Machine Learning

Weekly Progress

Name: Yaser Imtiyaz Ahmed Shaikh

Roll Number: 2060

Iymca

Week 2

Goa Business School

What is Machine Learning?

- Identifies relevant data sets and prepares—them for analysis
- Chooses the type of ML Algorithm then to use.
- Builds on analytical model based on the chosen algorithm.
- Train the model on test data sets, Revising as it is needed.
- Runs the model to generate Scores and other findings.

General Description:

Machine learning (ML) is a type of artificial Intelligence (AI) allows Software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning is a tool for tuning Information Into knowledge. Machine learning concerned with the use of data & algorithms that enable machines to imitate human learning so that they are capable of performing some sort of predictions by learning from Input examples. There are multiple forms of learning Machine learning approaches

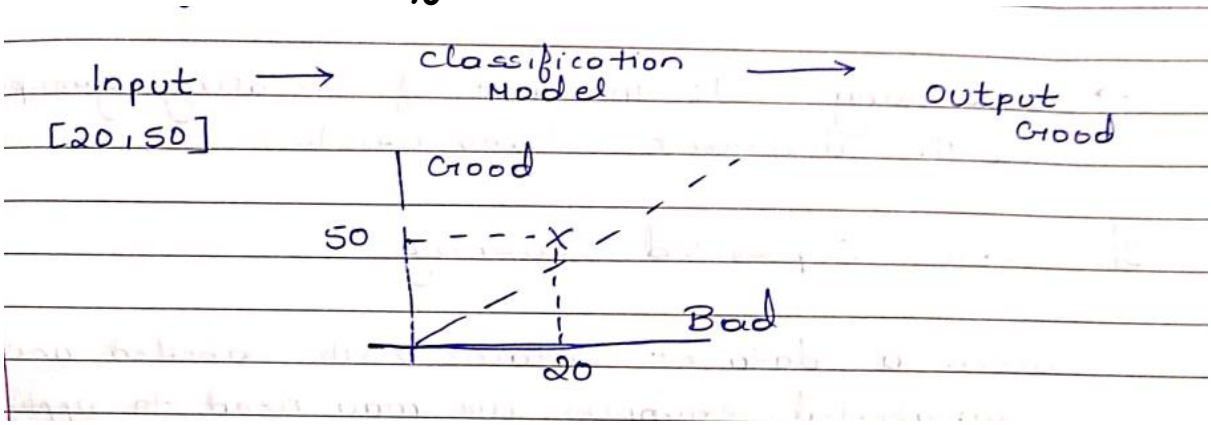
Supervised learning:

In Supervised learning the goal learns the mapping of the rules) between a set of Inputs and outputs: The dataset Includes labelled .
For example the Inputs could be the weather forecast, and the outputs would be the visitors to the beach. The goal in supervised learning would be to learn the mapping that describes the relationship between temperature and number of beach Visitors.
The output from a supervised machine learning model could be a category from a finite set. Example [low, med, high] for the number of visitors to the beach.

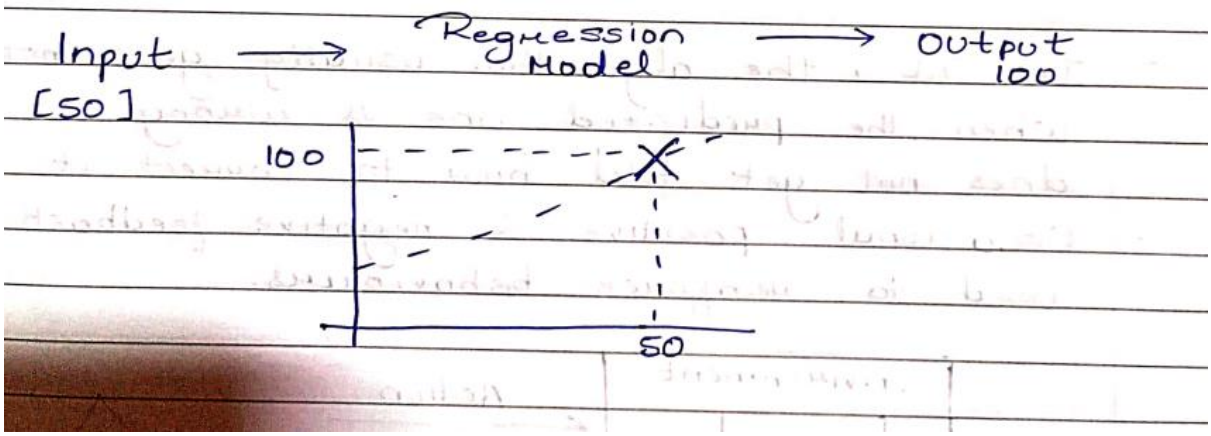
Input (temperature = 20) → model → output =
[visitors = high]. When this is case, it's is deciding how Input and so is known as classification.

Input [temperature = 300) ➔ model ➔ outputs = [Visitors2 = 200) when this is the case it is known as Regression.

A] Classification:- used to group the similar data points into different sections in order to classify them.



B] regression: regression outputs a number rather than a class.



Unsupervised learning:

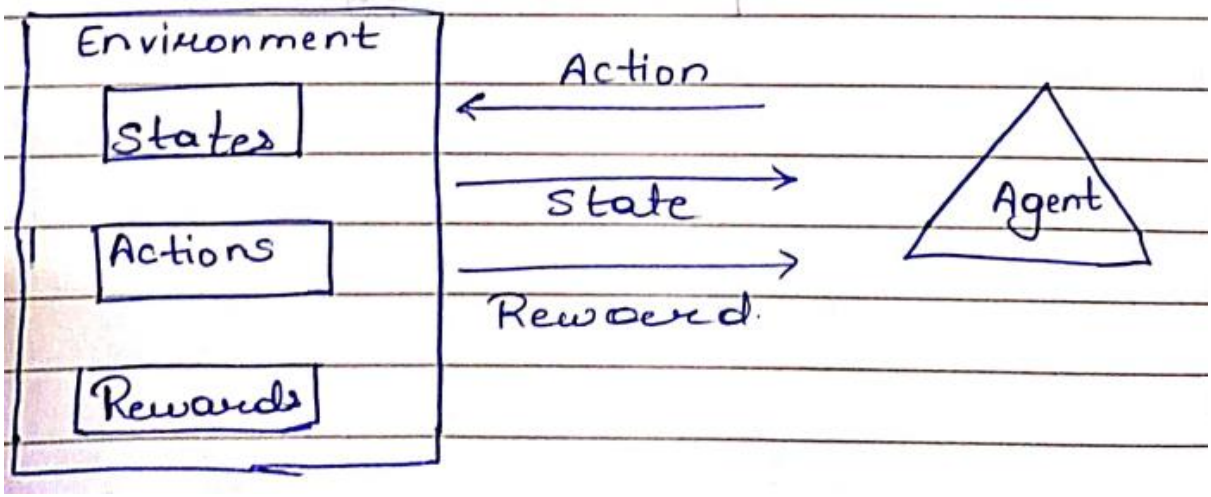
In Unsupervised learning only Inputs is provided there are no labelled examples outputs to aim for useful when we want perform clustering dimensionality reduction or outlier detection clustering is the act of creating groups with different characteristics.

Semi-Supervised learning:

When a dataset contains both labelled and unlabelled examples, we may need to apply a Semi - Supervised learning algorithm.

Reinforcement Learning:

In RL, the algorithm usually gets Instructed When the predicted ans is wrong but it does not get told how to correct it. Occasional positive & negative feedback is used to Reinforce behaviours.



Machine learning Applications

Recommendation Engines
Speech Recognition
Computer Vising
Automated trading s
Customer Service.

Quote

Celebration should not be limited to a particular occasion. Your whole life, your very existence should become a celebration.
Teach us from whom we can learn AI Ibrilliantly.

- Patrick Henny Winston
- Jom mitchel
- Andrew ng
- Bishop

1. Interpolation v/s Extrapolation
2. convex v/s concave
3. Deep learning
4. Curse of dimensionality
5. long term Aim and short term aim of AI
6. AGI vs Narrow Al

Narrow AI	General AI
• Interior to human Intelligence • Also known as weak AI • Lacks artificial Consciousness or cognitive abilities	• Similar to human Intelligence • Also known as strong AI or full AI • Has human-like consciousness and cognitive abilities. • Yet to become reality

• IBM Watson, Aplhago and Google Assistant are some example. • Can't solve unfamiliar problems	• Can solve unfamiliar problems.

➔ Advantage of Neural Networks – Benefits Deep Learning.

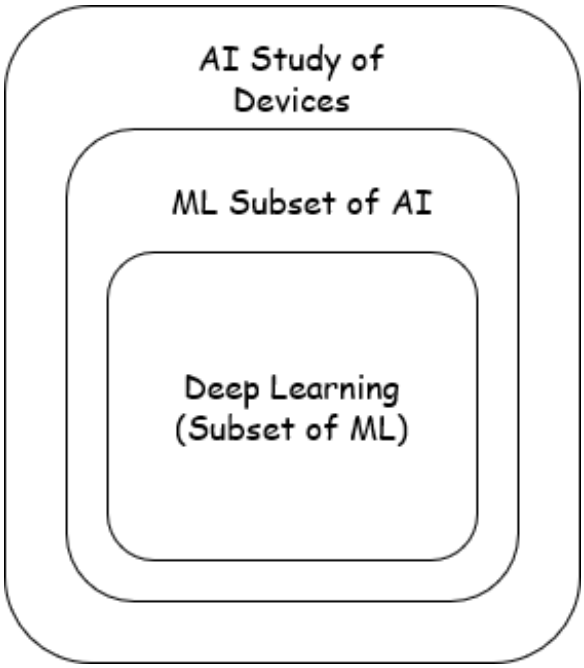
Artificial Intelligence - Past, Present and future

➔ Past

Machine learning -

Cannot generalize ie Connot be done using ML

- Data driven approach
- Compute power cpu
- lot of data
- Algorithm
- Revolutionaries



Neural Network and AI : Underlying Assumption

There are principles giving rise to Intelligence (machine, human or animal) via learning, Simple enough that they can be described Compactly, Similarly to the laws of physics ie our Intelligence is not just the result of a huge bag of tricks e pieces of Knowledge, but of general mechanisms to acquire knowledge.

Grand Engineering challenges for the 21st century

The 14 Ground challenges are :

- Advance Personalized learning
- Make Solar Energy Economical
- Enhance Virtual Reality
- Reverse Engineer the brain
- Engineer better medicines
- Advance health Informatics
- Restore and Improve urban Infrastructure
- Secure Cyberspace
- Provide access to clean water
- Provide Energy from fusion
- Prevent nuclear terron
- Manage the nitrogen cycle
- Develop carbon sequestration Method
- Engineer the tools of Scientific Discovery

Explainability

Aren't Neural n/w a black box? not really

Transformer model

Step 1: Pre train a model on this "fill in the blanks using large- amounts of self-supervised text:

Step 2: Fine - tune the model on Individual language toutes with small amounts of data.

The five tribes of Machine learning.

Where Does Knowledge come from?

- Evolution E
- Experience
- Culture
- Computers

So how do Computers Discover new knowledge?

- Fill in gaps in existing knowledge
- Emulate the brain
- Simulate evolution
- Systematically reduce Uncertainty
- Notice similarities bet old & new

The 5 tubes of ML

Tube	Origin	Master Algorithm
• Symbolists • Connectionists • Evolutionaries • Bayesians • Analogizers	• Logic, Philosophy • Neuroscience • Evolutionary Biology • Statistics • Psychology	• Inverse deduction • Back Propagation • Genetic Programming • Probabilistic Inference • Kernel Machines.

Probabilistic Inference

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Likelihood	Prior
How probable is the evidence given that our hypothesis is two?	How probable was our hypothesis before observing the evidence?

$$P(H|e) = \frac{P(e|H)\,P(H)}{P(e)}$$

Posterior	Marginal
How probable is our hypothesis given the observed evidence?	How probable is the new evidence under all possible hypothesis? $P(e) = \sum P(e|Hi)\,p(Hi)$

➔ Representation:

- Probabilistic logic (example: Markov logic networks)
- Weighted formulas ➔ Distribution over states

➔ Evaluation

- Posterior Probability
- User-defined objective function

➔ Optimization

- Formula discovery Genetic programming
- Weighted learning_: Back propagation

Curse of dimensionality

The curse of dimensionality refers to various phenomena that arise when analyzing and organizing data in high-dimensional spaces that do not occur in low-dimensional settings such as the three-dimensional physical space of everyday experience.

- Give fractions number
- Generalization
- Price of learning collections

Ml apps : machine learning operations

Machine learning operations (MLOps) is the use of machine learning models by development/operations (DevOps) teams. MLOps seeks to add discipline to the development and deployment of machine learning models by defining processes to make ML development more reliable and productive.

Data Drift: independent variables changes

Change in model input data that leads to model performance degradation

Concept Drift: dependent variable changes

Statistical properties of the target variable, which the model is trying to predict, change over time in unforeseen ways

What are the chances done in past few years?

Ans: exponential changes.

What is version space algorithm?

The version space search algorithm (Mitchell,1982) is a form of inductive learning implemented as search through a concept space. In our discussion, we use PROLOG syntax to represent concepts. So color(ball, red) represents the fact that the color of the ball is red.

When we are hired in the top company what skills You must have?

- Interpersonal skills. ...
- Learning/adaptability skills. ...
- Self-management skills. ...
- Organizational skills. ...
- Computer skills. ...
- Problem-solving skills. ...
- Open-mindedness. ...
- Strong work ethic.

Supervised Learning

Supervised learning is an approach to creating artificial intelligence (AI), where a computer algorithm is trained on input data that has been labelled for a particular output.

Self- Supervised Learning

Self-supervised learning (SSL) is a method of machine learning. It learns from un-labelled sample data. It can be regarded as an intermediate form between supervised and unsupervised learning. It is based on an artificial neural network.

Unsupervised Learning

The use of artificial intelligence (AI) algorithms to identify patterns in data sets containing data points that are neither classified nor labelled.

Semi Supervised learning

Combination of supervised and unsupervised learning. It uses a small amount of labelled data and a large amount of un-labelled data, which provides the benefits of both unsupervised and supervised learning while avoiding the challenges of finding a large amount of labelled data.

Finding approximation Function

X -> Y

Size -> Price

f(X) -> Y

No free lance

Person who pursues a profession without a long-term commitment to any one employer.

Literature

Which is intended or understood.

Classification

the act or process of classifying.

Systematic arrangement in groups or categories according to established criteria specifically : taxonomy.

➔ class, category

Write a program so that understand the pixel of image and output.

Reference Link To Google Collab : Click on Image

Reinforcement Learning

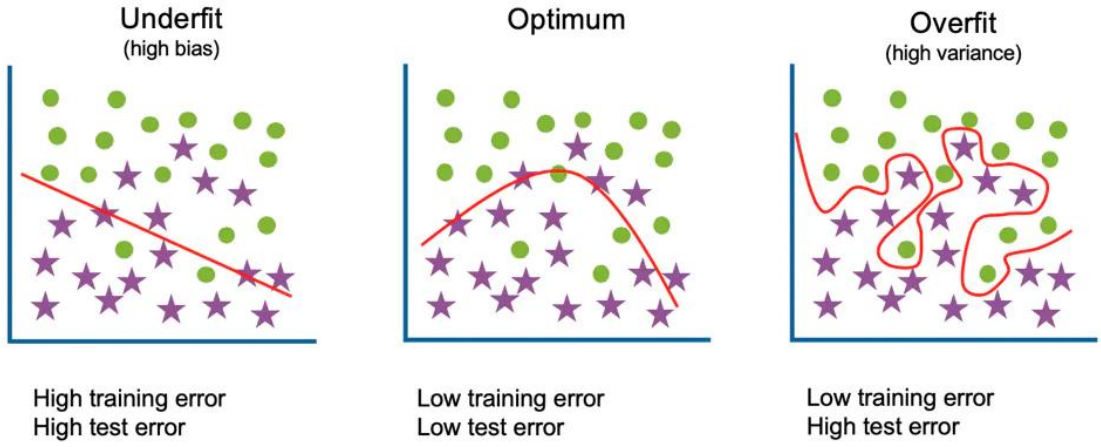
Reinforcement learning is a machine learning training method based on rewarding desired behaviours and/or punishing undesired ones.

Version Space

A version space is a hierarchical representation of knowledge that enables you to keep track of all the useful information supplied by a sequence of learning examples without remembering any of the examples.

Underfitting *is a scenario in data science where a data model is unable to capture the relationship between the input and output variables accurately, generating a high error rate on both the training set and unseen data.*

Overfitting *is a concept in data science, which occurs when a statistical model fits exactly against its training data. When this happens, the algorithm unfortunately cannot perform accurately against unseen data, defeating its purpose.*



Page No.:

Date:

04/07/2022

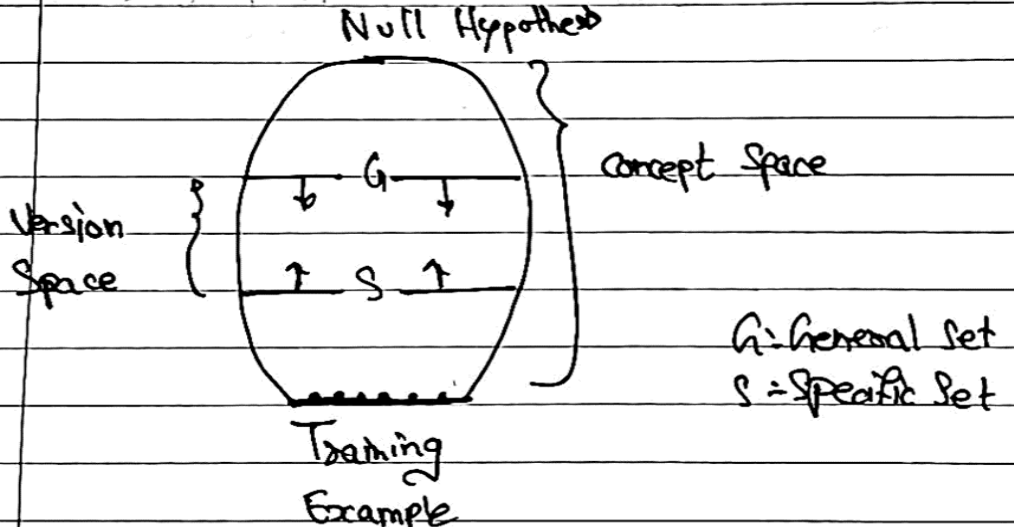
Representation lang for Cars

→ Origin : {japan, USA, britain,
→ Manufacture: {honda,
→ Color: {blue, green, red, white
→ Decade: {1988, 1960, 1970, 1990, 2000}
→ Type: {economy, luxuary, sports}

Concept Description

→ Origin : Japan
→ Manufacture = x2
→ Color: x3
→ Decade: x4
→ Type: Economy
→ The above descriptions for concepts for japanese economy car.

Version Space



Algorithm

→ A representation
in
→ Co description that is consistent

7/24

Solution:

Origin	Manufacturer	Color	Decade	Type	Example Type
Japan	Honda	Blue	1980	Economy	+ve
Japan	Tyota	Green	1970	Sports	-ve
Japan	Tyota	Blue	1990	Economy	+ve
USA	Chrysler	Red	1980	Economy	-ve
Japan	Honda	White	1980	Economy	+ve

Step 1:

→ Initialize G to a singleton set that includes everything.
→ Initialize S to a singleton set that

2) Negative Example: (Japan, Tyota, Green, 1970, Sports)
Specialize G to exclude the -ve example

G = {(?, Honda, ?, ?, ?),
(?, ?, Blue, ?, ?),
(?, ?, ?, 1980, ?),
(?, ?, ?, ?, Economy)}

S = {(Japan, Honda, Blue, 1980, Eco)}

- we G get specified
- we S get generalised

Repeat variable by variable

G_{En} to - Specific hypothesis

3) Positive Example: (Japan, Tyota, Blue, 1990, Economy)
Prune G to exclude description inconsistent with the +ve example

$G = \{(?,?,Blue,?,?), (?,?,?,?,Eco)\}$

$S = \{(Japan, ?, Blue, ?, Eco)\}$

4) Negative Example : (USA, Chrysler, Red, 1980, Economy)

Specialize G to exclude the -ve example

$G = \{(?,?\ Blue,?), (Japan, ?,?,?, Eco)\}$

$S = \{(Japan, ?, Blue, ?, Eco)\}$

5) Positive Example. (Japan, Honda, White, 1980, Economy)
Prune G to exclude descriptions inconsistent with the example.
Generalise S to include the example.

$G = \{(Japan, ?, ?, ?, Eco)\}$
$S = \{(Japan, ?, ?, ?, Economy)\}$

G and S are singleton set $S = G$
Converged
No more data, Algorithm stops.

4