

Machine Learning

Assignment I

Dhananjay Purohit

EN19CS30110

CS-II-B

- 1 Define machine learning. Explain different perspectives & issues in machine learning.

Ans Machine Learning is an application of artificial Intelligence (AI) that provides systems the ability to automatically learn & improve from experience without being explicitly programmed.

Machine learning focuses on the development of computer programs that can access data & use it to learn for themselves.

Issues in Machine Learning -

It provides businesses with the knowledge to make more informed, data-driven decisions that are faster than traditional approaches.

- a) Understanding which processes need automation -
It is becoming increasingly difficult to separate fact from fiction in terms of machine learning today.
The easiest processes to automate are the ones that are done manually every day with no variable output.
- b) Lack of Quality Data -
While enhancing algorithms often consumes most of the time of developers in AI, data quality is essential for the algorithms to function as intended.
- c) Inadequate Infrastructure -
ML requires vast amounts of data churning capabilities. Legacy systems often can't handle the workload & buckle under pressure. You should check if your infrastructure can handle ML. If it can't, you should look to upgrade complete with hardware acceleration & flexible storage.
- d) Lack of skilled Resources -
Deep analytics & ML is their current forms and still new technologies. Thus, there is a shortage.

of skilled employees available to manage or develop analytical content for ML. Data Scientists often need a combination of domain experiences, as well as in-depth knowledge of science, technology, & mathematics.

- 2 Explains the steps involved in designing a learning system?

Ans Steps for designing learning systems are -

Step 1 Choosing the training Experience:

The very important first task is to choose the training data or training experience which will be fed to the ML algorithm.

The training experience will be able to provide direct or indirect feedback regarding choices. Second important attribute is the degree to which the learner will control the sequences of training examples.

Step 2 Choose target function:

The next important step is choosing the target fun'. It means according to the knowledge

fed to the algorithm the machine learning will choose next move function which will describe what type of legal moves should be taken.

Step 3 choosing representation for target function:

When the machine algorithm will know all the possible legal moves the next step is to choose the optimized move using any representation i.e. using linear equations, hierarchical graph representation, tabular form etc. The next move function will move the Target move like out of these moves which will provide more success rate

Step 4 choosing function Approximation Algorithm:

An optimized move cannot be chosen just with the training data. The training data had to go through with set of example & through these examples the training data will approximates which steps are chosen and after that machine will provide feedback on it.

Step 5 Final Design:

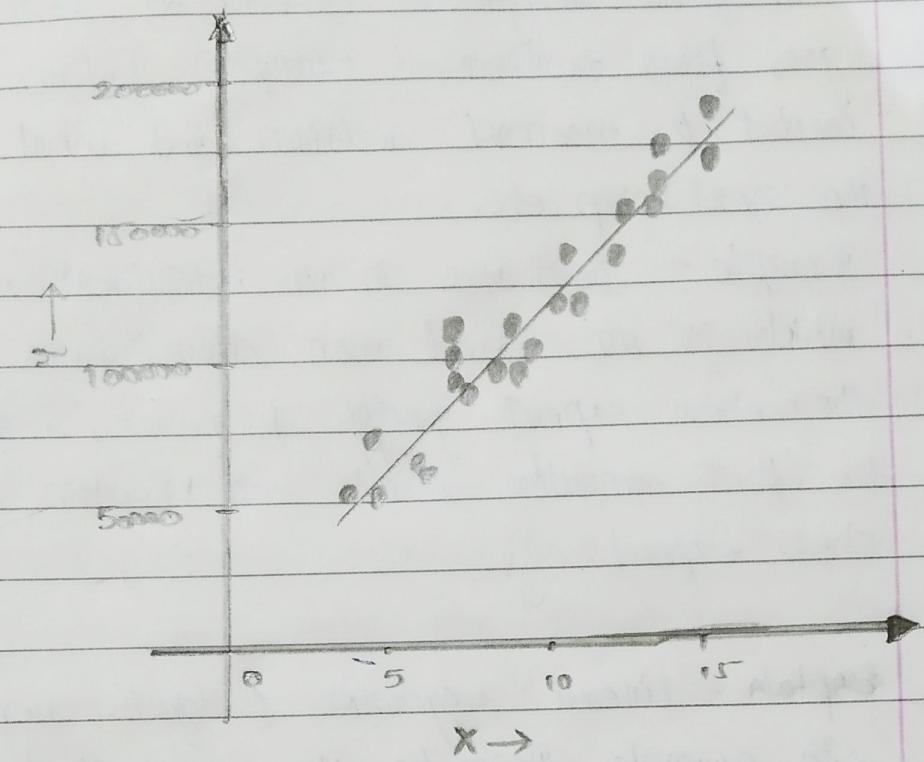
The final design is created at last when system goes from number of examples, failures & success, correct & incorrect decision and what will be the next step, etc.

Example - Deep blue is an intelligent computer which is ml based won chess game against the chess expert garry kaspakov, and it become the first computer which had beaten a human chess expert.

- 3 Explain linear regression & logistic regression model with example. Describe the application areas for both the models.

Ans • Linear Regression Model -

Linear Regression is a machine learning algorithm based on supervised learning. It perform a regression task. Regression model a target prediction value based on independent variable. It is mostly used for finding out the relationship between variables & forecasting.



Example

It can be used to quantify the relative impacts of age, gender, & diet on height.

The table below shows some data from the early days of the Italian clothing company Benetton. If we use advertising as the predictor variable, linear regression estimates that

$$\text{Sales} = 108 + 23 \text{ Advertising}$$

Year	Sales	Advertising
1	651	23
2	762	26
3	856	30
4	1063	34
5	1190	43
6	1298	48
7	1421	52
8	1640	57
9	1518	58

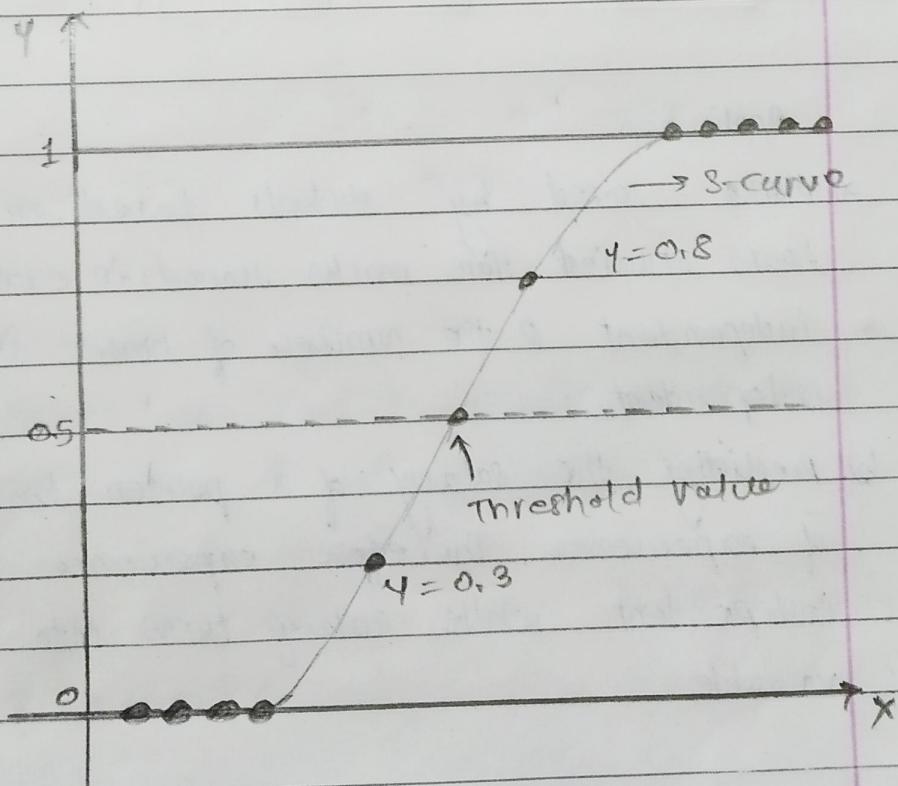
Application

- Marks scored by students based on number of hours studied. Here, marks scored in exams are independent & the number of hours studied is independent.
- Predicting the salary of a person based on years of experience. Therefore, experience becomes the independent while salary turns into the dependent variable.

- Logistics Regression Model -

Logistic regression is one of the most popular machine learning algorithms, which comes under the supervised learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables.

It can be either Yes or No, 0 or 1, true or false, etc. but instead of giving the exact value as 0 or 1, it gives the probabilistic values which lies between 0 & 1.



Example

There is a dataset given which contains the information of various user obtained from the social networking sites. There is a car making company that has recently launched a new SUV car. So, the company wanted to check how many users from the dataset, wants to purchase the car.

User ID	Age Gender	Gender Age	Estimated Salary	Pur.
15624510	19	M	19000	0
15610123	35	F	20000	0
15612567	26	M	43000	0
15313248	27	F	57000	0
15715213	19	F	76000	0
15810123	27	F	58000	0
15215260	27	M	84000	0
153112112	32	M	180000	1

Applications

Logistic regression is used in various fields, including machine learning, most medical fields, & social sciences.

- 4 Explain how gradient descent algorithm works to optimize cost function for linear regression?

And Gradient descent is the process of minimizing a function by following the gradient of the cost function.

The way this works is that each training instance is shown to the model one at a time. The model makes prediction for a training instance, the error is calculated & the model is updated in order to reduce the error for the next prediction.

$$w = w - \alpha * \text{delta}$$

where, w is the coefficient or weight being optimized.

α is learning rate

gradient is the error for the model on the training data attributed to the weight.

- 5 Explain k-nearest neighbour learning algorithm with example?

And K-Nearest Neighbour is one of the simpler ML algo. based on supervised learning technique.

K-NN algo. stores all the variables data & classified a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suited category by using K-NN algo. It is also called a lazy learner algo. because it does not learn from the training set.

Example -

Suppose, we have an image of a creature that looks similar to cat & dog, but we want to know either it is a cat or dog. So, for this identification, we can use the KNN algorithm, as it works on a similarity measure. Our KNN model will find the similar features of the most similar features it will put it in either cat or dog category.