1) Illustrate Principal Component Analysis in Machine Learning.

A) Principal Component Analysis:

* Principal Component Analysis is an unsupervised learning algorithm that is used for the dimensionality reduction in machine learning.

* It is a statistical process that converts the observations of correlated features into a set of linearly uncorrelated features with help of orthogonal transformation. These new transformed features are called "Principal Components".

Principal Components in Principal Component Analysis:

The number of these principal Components are either equal to or less than original features present in dataset.

Properties of these principal Components:

* the principal Component must be linear combination of the Original features.

* these components are orthogonal i.e, the correlation b/w a pair of variables is zero.

* The importance of each component decreases when going to 1 to n, it means the 1 PC has most impotance, & n PC will have least importance.

Steps for PCA algorithm:

1) Getting the dataset:

Firstly, we need to take input dataset and divide it into two subparts X & Y, where X is training set, Y is validation set.

2) Representing data into a structure:

Now represent dataset into a structure. Such as we will represent 2D matrix of independent variable X. Here each row corresponds to data items, and column corresponds to features.

3) Standardizing the data:

Now we will standardize our dataset. Such as in particular column, the features with high variance are more important compared to features with lower variance. If the imp of features is independent of variance of feature, then we will divide each data item in a column with S.D of column. Here we will name the matrix as "Z".

4) Calculating the Covariance of Z:

We will take matrix Z, and will transpose it. After transpose, we will multiply it by Z.

5) Calculating the eigen values & Eigen Vectors:

Eigen vectors are the directions of the axes with high information and the Coefficients of these eigenvectors are the eigen values.

6) Sorting eigen vectors:

Here we will take eigenvalues and sort them in decreasing order, and simultaneously sort eigenvectors accordingly in matrix P of eigen values. Resultant matrix named as P*.

7) Calculating Principal Components:

We will multiply P* matrix with Z. In resultant matrix Z*, each observation is linear combination of original feature.

8) Remove unimportant features from new dataset:

Here we will only keep important features in new dataset & unimportant features will be removed out.

Applications of Principal Component Analysis:

* PCA is mainly used as Dimensionality reduction technique in various AI applications such as computer vision, image compression etc;
* It can also be used for finding hidden patterns if data has high dimensions.
* Some fields where PCA is used are Finance, datamining, Psychology etc;

Scanned with CamScanner

2) Explain linear Discriminant Analysis for Machine Learning

A) Linear Discriminat Analysis (LDA):

* LDA is a supervised classification method that is used to create ML models. these models based on dimensionality Reduction are used in marketing predictive analysis & image recognition.

* there are two areas that LDA helps in discovering - the parameters that can be used to explain relationship b/w group & an object - The classification preceptor model that can help in seperating the groups.

Extensions to Linear Discriminant Analysis:

* LDA is considered one of the Simplest & most effective methods available for classification. We have few variations as well as extensions discussed below.

a) Regularized Discriminant Analysis (RDA):

RDA is used for bringing regularization into variance or covariance estimation. this is done to moderate the impact that variables have on LDA.

b) Quadratic discriminant Analysis (QDA):

In QDA, diff classes use their own variance estimate. In case the no of input variable is more than usual, every class uses its covariance estimate.

c) Flexible Discriminant Analysis(FDA):

FDA makes use of inputs with non-linear Combinations.

Eg: Splines

Applications of LDA:

1) Face recognition:

* In Computer version, face recognition is considered as one of the most popular applications.

* Face recognition is carried out by representing faces using Large amount of pixel values.

* LDA is used to trim down the no of features to prepare grounds for using classification method.

* The new dimensions are combinations of pixel values that are used to create a template.

2) Customer Identification:

* If you want to identify customers on basis of likelihood that they will buy a product, you can use LDA to collect customer features.

* You can identify & choose those features that describe the group of customers that are showing higher chances of buying a product.

3) Medical:

LDA can be used to put diseases into diff categories such as severe, mild, or moderate. These are several patient Parameters that will go into conducting this classification task. This allow doctors to define pace of treatment.

3) Illustrate Cross validation in Machine Learning.

A) Cross validation in Machine Learning:

* It is a technique for validating the model efficiency by training it on subset of input data and testing on previously unseen subset of input data.

Basic steps of cross Validation:

* Reserve a subset of dataset as validation set.
* Provide training to model using training dataset.
* Now, evaluate model perfomance using validation set. If the model performs well with validation set, perform further step, else check for issues.

Methods used for cross-Validation:

1) Validation Set Approach:

* We divide our input dataset into a training set & test or validation set in validation set approach. Both subsets are given 50% of dataset
* But it has one of big disadvantages that we are just using a 50% dataset to train our model, so the model may miss out to capture imp info of dataset.

2) Leave-P-Out cross-Validation:

* In this approach, the p datasets are left out of training data. It means, if there are total n datapoints in original input dataset, then n-p data points will be

used as training dataset & p datapoints as validation set. this complete process is repeated for all Samples, & average error is calculated to know effectiveness of model.

* Disadvantage is, it can be computationally difficult for large p.

3) Leave one out cross-validation:

* In this approach, for each learning set, only one datapoint is reserved, and the remaining dataset is used to train the model. For n samples, we get n diff training set & n test set.

Features:

* The bias is minimum as all datapoints are used.
* The process is executed for n times, hence execution time is high.

*

4) k-fold cross-Validation:

* this approach divides input dataset into k groups of samples of equal sizes. These samples are called folds.

steps for k-fold cross validation:

* split the input dataset into k groups
* For each group:
    * Take one group as reverse or test dataset.
    * Use remaining groups as training dataset
    * Fit the model on training set & evaluate performance of model using test set

5) Stratified k-fold cross-Validation:

* This approach works on stratification concept, it is a process of rearranging the data to ensure that each fold is a good representative of complete dataset.
* To deal with bias & variance, it is one of the best approaches.

Limitations of Cross-Validation:

* For the ideal conditions, it provides optimum output. But for inconsistent data, it may produce a drastic result. So it is one of big disadvantages of cross-Validation.
* In predictive modeling, the data evolves over a period, due to which, it may face diff b/w training set & validation sets. Such as if we create a model for prediction of stock market values, & data is trained on previous 5 years stock values, but realistic feature values for next 5 years may drastically differ, so it is difficult to expect correct o/p for such situations.

Applications of Cross-Validation:

* This technique can be used to compare performance of diff predictive modeling methods.
* It has great scope in medical research field.
* It can also be used for meta analysis, as it is already being used by data scientists in field of medical statistics.

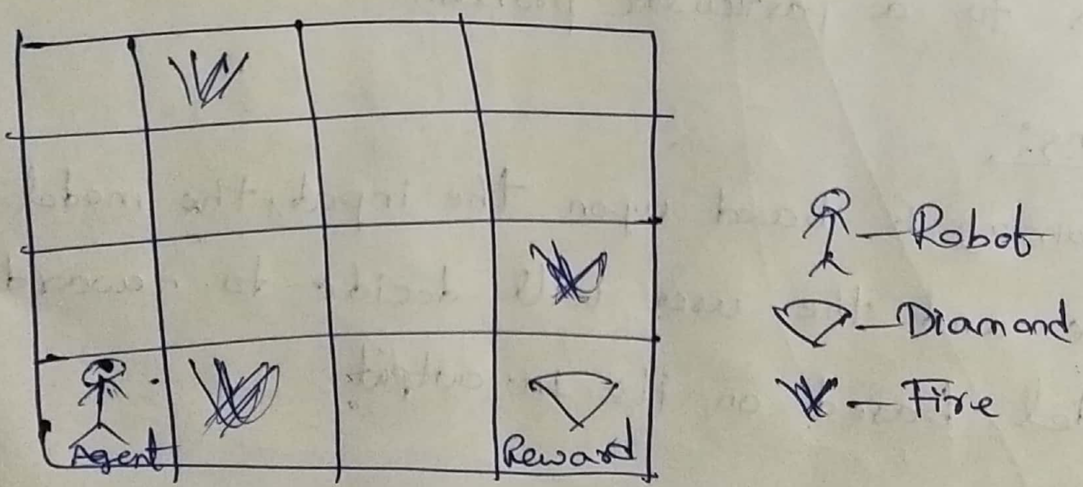
4) Demonstrate reinforcement learning with an example.

A) *Reinforcement Learning is an area of Machine Learning. It is about talking suitable action to maximize reward in a particular situation.

* In reinforcement Learning, there is no answer but reinforcement agent decides what to do to perform given task. In the absence of training dataset, it is bound to Learn from its experience.

Example:

We have agent and a reward, with many hurdles in b/w. the agent is supposed to find the best possible path to reach the reward.



* The above image shows the robot, diamond & fire. The goal of the robot is to get the reward that is the diamond and avoid the hurdles that are fired.

* The robot Learns by trying all the possible paths & then choosing the path which gives him the reward with Least hurdles.

* Each right step will give the robot a reward & each wrong step will ~~~~ subtract the reward of the robot.

* The total reward will be calculated when it reaches final reward that is the diamond.

Main points in Reinforcement Learning:

1) Input:

The input should be an initial state from which the model will start

2) Output:

There are many possible outputs as there are variety of solutions to a particular problem.

3) Training:

The training is based upon the input, the model will return a state and the user will decide to reward or punish the model based on its ~~in~~ output

4) The model keeps continues to learn.

5) The best solution is decided based on the maximum reward.

Applications of Reinforcement Learning:

* RL Can be used in robotics for industrial automation.
* RL Can be used in machine Learning & data processing.
* RL Can be used to create training Systems that provide custom instruction and materials according to the requirement of students.

RL Can be used in Large environments in following situations:

1) A model of environment is known, but an analytic solution is not available.
2) Only a simulation model of environment is given (the subject of simulation - based optimization)
3) The only way to collect information about the environment is to interact with it.

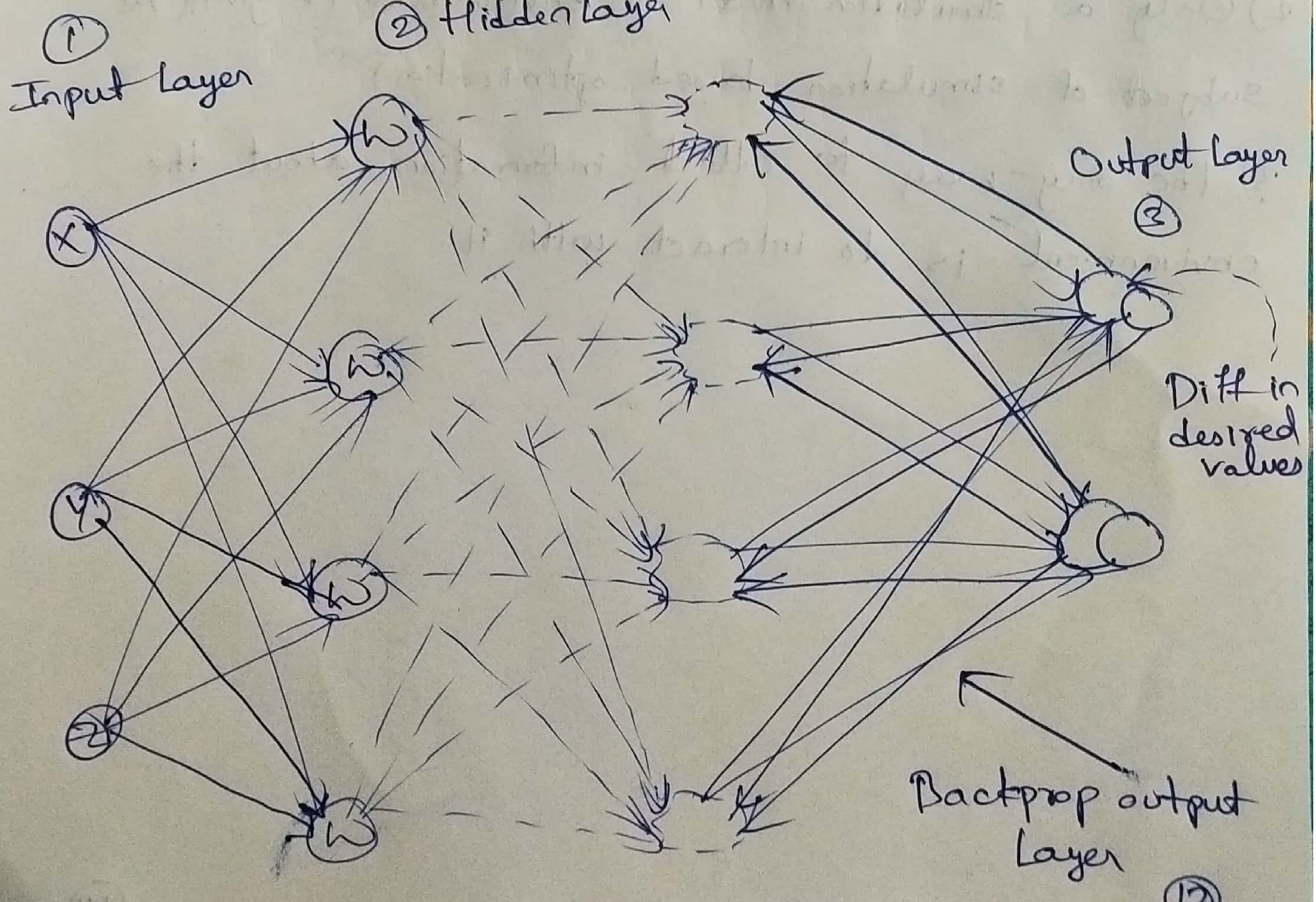
5) Explain Backpropagation algorithm.

A) Back Propagation:

* Back Propagation is the method of fine-tuning the weights of neural network based on error rate obtained in previous iteration.
* Proper tuning of weights allows you to reduce error rates & make model reliable by increasing its generalisation.

Working of BackPropagation Algorithm:

* BackPropagation algorithm computes gradient of loss function for a single weight by chain rule. It efficiently computes one layer at a time, unlike a native direct computation.



1) Inputs X, arrive through preconnected path

2) Input is modeled using real weights W. The weights are usually randomly selected.

3) Calculate output for every neuron from input layer, to hidden layers to output layer.

4) Calculate error in outputs.

Error$_B$ = Actual O/P − Desired O/P

5) Travel back from O/P layer to hidden layer to adjust weights such that error is decreased

Keep repeating process until desired O/P is achieved.

Types of Back Propagation networks:

1) Static Back Propagation:

It is one kind of Back propagation network which ~~provides~~ produces a mapping of static input for static O/P. It is useful to solve static classification issues like optical character recognition.

2) Recurrent Backpropagation:

Recurrent Backpropagation in data mining is fed forward until a fixed value is achieved. After that, the error is computed & propagated backward.

Key points of Back Propagation:

* Simplifies the network structure by elements weighted links that have least effect on trained network.
* It helps to assess the impact that a given input variable has on a network o/p. The knowledge gained from this analysis should be represented in rules.
* It is ~~usually~~ especially useful for deep neural networks working on error-prone projects, such as Image or speech recognition.

Disadvantages of using Back Propagation:

* Actual performance of Back propagation on a specific problem is dependent on i/p data.
* Back Propagation algorithm in data mining can be quite sensitive to noisy data.
* We need to use matrix-based approach for Back propagation instead of mini-batch.