Gradient Descent Algorithm

* Gradient Descent is a popular optimization algorithm used to minimize the cost function ~~by~~ in Machine Learning model.
* Gradient Descent will reduce the error by updating the weights which refers to the distance between the actual point and the hypothesis point.

Steps in Gradient Descent Alg:

Step 1: Initialize the model's parameter with some values

Step 2: Calculate ~~update~~ the cost function for the current values of parameters.

Step 3: Calculate the gradient of cost function for each value of the parameter.

Step 4: Update the values of the parameter by subtracting a small fraction of gradient. This small fraction is called learning rate.

Step 5: Repeat step 3 & 4 until the cost function converges to a minimum value.

Back Propagation Algorithm

* Back propagation is a technique which involves deviation of weights obtained from the summation function and activation function which is returned to the neural network which helps in adjusting the weights thus reducing the error.

Steps:

Step 1: FORWARD PASS: Input the input data to the network and compute the output.

Step 2: COMPUTE ERROR: Compare the predicted output with the true output and compute the error.

Step 3: BACKWARD PASS: Compute the gradient of the cost function of the network ~~to minimize the cost function~~ with the weights of the network.

Step 4: UPDATE THE WEIGHTS: Use the gradients to update the weights of the network to minimize the cost function.

SVD Algorithm

* Singular value Decomposition.
* SVD is a matrix factorization technique which decomposes a given matrix into three matrices which can be used for various tasks such as data compression, data visualization and noise reduction.
* It can be used in collabractive filtering and image processing.

Steps:

Step 1: compute $A * A^T$: Multiply the given matrix A with its transpose (A^T) to obtain a matrix of dimension $m \times n$.

Step 2: compute eigen values & vectors for $A * A^T$: compute the eigenvectors & eigen values for matrix obtained in step 1. The eigenvectors form the columns of matrix U and the square root of eigen values form the diagonal elements of matrix E.

Step 3: compute $A^T * A$: Multiply the ~~given ma~~ matrix A^T with the given matrix A to obtain a matrix of dimension $m \times n$.

Step 4: compute eigen values & vectors of $A^T * A$: Compute the eigen vectors & eigen values of the matrix obtained in step 3. The eigen vectors form the column of matrix V.

Step 5: construct the SVD Decomposition of matrix A.

Multiply the matrices obtained in step 2, 3, 4 SVD can be used to construct the decomposition of matrix A.

$$A = U \times E \times V^T$$

Dimensionality reduction:

* Dimensionality reduction is a technique used in data analysis and machine learning to reduce the number of features or variables while retaining the most relevant information. This is used to address the problem of "curse of dimensionality" which refers to the difficulty of working with high dimensional data which requires high memory and computation.

Types of Dimensionality Reduction techniques:

* Feature selection: In this method, a subset of original features is selected which is most relevant to the problem at hand. This can be done by various criteria such as correlation, mutual information, statistical tests etc. Then the selected features can be used for further analysis or modeling.
* Feature extraction: In this method, we transform the original features into a new set of features to capture the most important information in the data. This can be done by using PCA, LDA and SVD techniques. Then the new feature can be used for further analysis and modeling.

Dimensionality Reduction techniques.:

1) PCA :

* Principal component Analysis
* Unsupervised machine learning technique
* Here the original features will be transformed to new features to capture the most important information in the data.
* This can done by finding the principal component ~~by~~ from the data, which are the directions in which the data varies the most.
* Can be used for data compression and data visualization.

2) LDA :

* Linear Discriminant Analysis.
* Supervised machine learning technique.
* LDA follows feature extraction technique
* ~~LDA aims at minimizing the differences between 2~~
* LDA aims at maximizing the separation between different classes and minimizing the variance within each class

3) SVD :

* Singular value Decomposition
* Follows feature extraction technique
* SVD is a matrix factorization technique in which the given matrix will be decomposed into 3 matrices which can be used in various tasks such as data compression, data visualization and noise reduction.
* It can be used in collaborative filtering and image processing.