1.What's Dimensionality Reduction?

**Solution: -**

* Basically, independent variables called features. The higher the number of features, the harder it gets to visualize the training set and then work on it. Sometimes, most of these features are correlated, and hence redundant. This is where dimensionality reduction algorithms to play.
* Dimensionality reduction is the process of reducing the number of random variables under consideration, by obtaining a set of principal variables. It can be divided into feature selection and feature extraction.
* A classification problem that relies on both humidity and rainfall can be merged into just one underlying feature, since both are correlated to a high degree. Hence, we can reduce the number of features in such problems.

There are two components of dimensionality reduction:

* **Feature selection:** In this, we try to find a subset of the original set of variables, or features, to get a smaller subset which can be used to model the problem. It usually involves three ways:
  + Filter
  + Wrapper
  + Embedded
* **Feature extraction**: This reduces the data in a high dimensional space to a lower dimension space, i.e. a space with lesser no. of dimensions.

**Methods of Dimensionality Reduction**

The various methods used for dimensionality reduction include:

* Principal Component Analysis (PCA)
* Linear Discriminant Analysis (LDA)
* Generalized Discriminant Analysis (GDA)

Dimensionality reduction may be both linear or non-linear, depending upon the method used. The prime **linear method**, called Principal Component Analysis, or PCA.

It works on a condition that while the data in a higher dimensional space is mapped to data in a lower dimension space, the variance of the data in the lower dimensional space should be maximum

Diagram

Description automatically generated

It involves the following steps:

* Construct the covariance matrix of the data.
* Compute the eigenvectors of this matrix.
* Eigenvectors corresponding to the largest eigenvalues are used to reconstruct a large fraction of variance of the original data.

we are left with a lesser number of eigenvectors, and there might have been some data loss in the process. But the most important variances should be retained by the remaining eigenvectors.

**Advantages of Dimensionality Reduction**

* It helps in data compression, and hence reduced storage space.
* It reduces computation time.
* It also helps remove redundant features, if any.

**Disadvantages of Dimensionality Reduction**

* It may lead to some amount of data loss.
* PCA tends to find linear correlations between variables, which is sometimes undesirable.
* PCA fails in cases where mean and covariance are not enough to define datasets.
* We may not know how many principal components to keep- in practice, some thumb rules are applied.

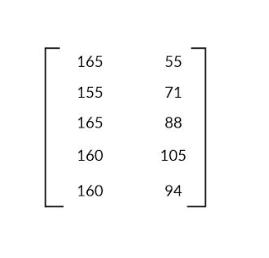
PCA is a statistical procedure to convert observations of possibly correlated variables to ‘principal components’ such that:

* They are uncorrelated with each other.
* They are linear combinations of the original variables.
* They help in capturing maximum information in the data set.

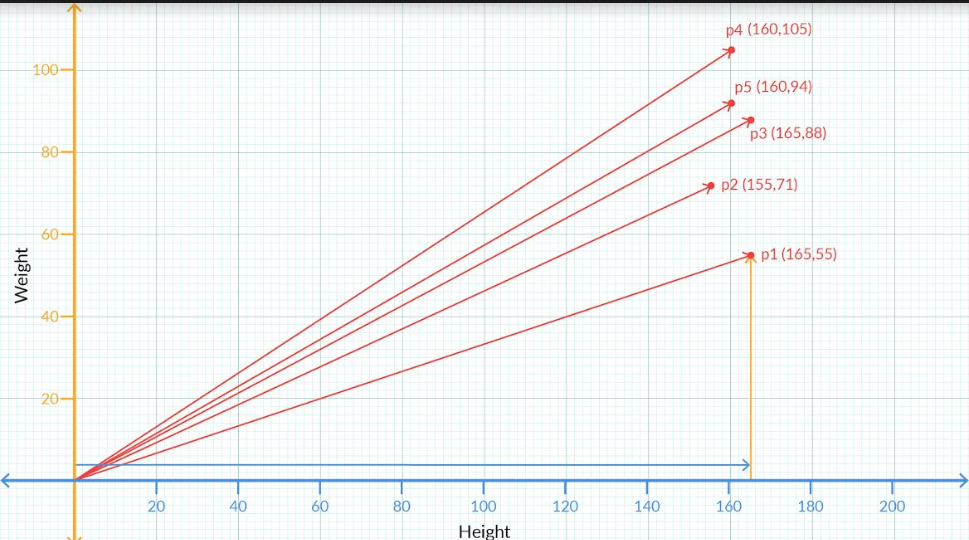
PCA are basically: -

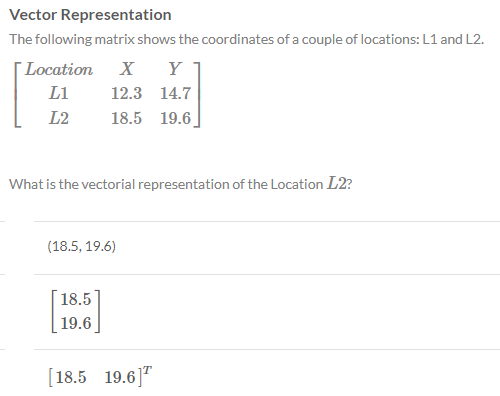
* It is an unsupervised technique.
* Principal components are the linear combinations of original variables
* Principal components are constructed to capture the maximum information



The height and weight information can be represented in the form of a matrix as follows:-

with each row representing a particular patient's data and each column representing the original variable. Geometrically, these patients can be represented as shown in the following image:





**1. Vectors have a direction and magnitude**

Each vector has a direction and magnitude associated with it. The direction is given by an arrow starting from the origin and pointing towards the vector's position. The magnitude is given by taking a sum of squares of all the coordinates of that vector and then taking its square root.

For example, the vector (2,3) has the direction given by the arrow joining (0,0) and (2,3) pointing towards (2,3). Its **magnitude** is given by √22+√32=√13.

Similarly, for a vector in 3 dimensions, say (2,-3,4) its direction is given by the arrow joining **(0,0,0) and (2,-3,4) pointing towards (2,-3,4)**. And as in the 2D case, we get the magnitude of this vector as √(2)2+√ (−3)2+√ (4)2=√29 .

2. **Vector Addition**

When you add two or more vectors, we essentially add their corresponding values element-wise. The first element of both the vectors get added, the second element of the both get added and so on.

For example, if you've two vectors say

V1=(2,3) and V2=(1,2) then

V1+V2=(2+1,3+2)=(3,5).

3. In the i, j notations that we introduced earlier, the above addition can be written as V1+V2=(2i+3j)+(i+2j)=(2+1)i+(3+2)j=3i+5j

Similarly, this idea can be extended to multiple dimensions as well.

4. Scalar Multiplication

If you multiply any real number or scalar by a vector, then there is a change in the magnitude of the vector and the direction remains same or turns completely opposite depending on whether the value is positive or negative respectively.

**Example of PCA:-**

