Al-->AI helps the machine to think without the human intervention(the machine will take its own decision)

ML-->it will focus on the algorithm which helps computer to learn the data(like predicting datas)

The inputs are called as **feature vector**.

It has two type:

1.qualitative data(nominal data,ordinal data)--> categorical data

2.quantitative data(discrete ,continuous)--->numerical data

1.**qualitative**:

There will be a finite categories or groups.

example: a boy ,a girl,we can say like two gender or categories,

\***Nominal data**:(no inherent data)

There is not any inherent order built into either of these categorical dataset.

**one hot encoding**: its used to convert the categorical data into binary data(0,1)

\***Ordinal data**:(inherent order)

ex: education scale

The categorical data that has a clear order or ranking order among its categories, but the difference between those categories are measurable and consistent.

2.quantitative:

There will be numerical values data.

: numerical value of data

it only take specific values

the length of something (ex the temperature degree, the length(also count)

\***continuous:**

it can take on an infinite number of values within a specific range (e.g., between 1.0 and 2.0, you can have 1.1, 1.11, 1.111, etc.).

\*it can be measured.we cant count them.

\*it can take any values(0.1,22.3)

**Type of prediction:**

classification--predict discrete classes

**1.binary classification:**

two prediction is or not (0,1)

**2.multi class:**

predict more than two or more categories

**K-Nearest Neighbor:**

* K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories.
* Take the label which has a majority to its nearest.
* To find the our distance function we use Euclidean distance.

**Precision and recall:**

* Precision measures the proportion of correct positive predictions out of all positive predictions made by the model.

**Recall:**

* Recall measures the proportion of correct positive predictions out of all actual positive cases in the dataset.
* F1-score is the combination of the recall and precision.

**Naïve bayes:**

The occurrence of a certain feature is independent of the occurrence of other feature.

it depends on the principle of [Bayes' Theorem](https://www.javatpoint.com/bayes-theorem-in-artifical-intelligence).

* Posterior probability(the probability that a person has the disease is (0.01) because it's rare)
* Likelihood (If a person **has** the disease the test)
* Prior probability (If a person **does not have** the disease, but tested as positive)
* Evidence (The overall probability of getting a positive test result)

**Logistic regression:**

where the goal is to predict the probability that an instance belongs to a given class or not.

Trying to fit our data to the sigmoid function

\*When we have one datapoint or one feature x then it called simple logistic regression(x0=simple LR)

\* when we have multiple features to built a model is called multiple LR(X0,X1……Xn=multiple LR)

**SVM(support vector machine):**

SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future.

**Hyperlane:**

We are separating the two classes using the hyperlane(A decision boundary that separates different classes in the feature space.)

**Kernel trick:**

If the data is not linearly separable in the original space, SVM uses a kernel function to map the data into a higher-dimensional space where it becomes separable.

**Neural network:**

 neurons interconnected to one another, artificial neural networks also have neurons that are interconnected to one another in various layers of the networks.

Each of the hidden layer has the neuron all the input has a weight ed by some values and then they get multiplies and the sum of all these goes into the neuron .

**Activation function:**

Sigmoid function: its run between zero and one.

Tanh function: between negative and all the way to the one.

Relu : anything less than zero.

**TENSERFLOW:**

A opens ource library to hep you to develop and train ml models.

**plot\_history()-**visualize the training history of a model in TensorFlow/Keras.

The history parameter is an object returned by the model.fit() method which contain the loss and accuracy.

**fig, (ax1, ax2)** = plt.subplots(1, 2, figsize=(10, 4))-figure size.

**Epoch**= the number of training iterations

**Binary crossentropy** = commonly used for binary classification tasks

When you pass a history object to this function, it will generate two side-by-side plots:

1. **Loss vs. Epochs** (Left):
   * Shows how the training and validation losses evolve with epochs.
   * Ideally, both losses decrease and plateau at some point.
2. **Accuracy vs. Epochs** (Right):
   * Shows how the training and validation accuracies evolve.
   * Ideally, both accuracies increase and plateau at some point.

This visualization is particularly useful for diagnosing and improving model performance during training.

**Linear regression:**

Linear regression we using some line of fit that will help to decrease this measure of error with respect to all the datapoints that we have in the datasets and try to come up with the best prediction for all of them this is known as simple linear regression.

**Linearity**:The relationship between independent and dependent variables is linear.

**Independence**: all the points in the dataset should be independent and shouldnot affect other

**Homoscedasticity**:The variance of residuals (errors) is constant across all levels of the independent variables.

**Normality**: Residuals(the errors) are normally distributed.

**Evaluating the linear regression:**

**Mean absolute error:**

It take all the error(residuals) and sum up the distance for all of them,take the average and find how far off they are.

**Mse: mean squared error:**

Sumup all the error and square the m all then divide them by n in order to find the mean,like taking all the different values and squaring them first before I add them to one and another and I divide them. Helps to punish the large error.

**Root Mean Squared Error (RMSE)**

Taking the square root ensures the error is in the same units as the original values, making it easier to interpret.

**R2-co eff of determination**:

RSS=sum of squared residuals

TSS=total sum of squares