Model is like the brain of the dataset (sort of) you can also consider it to be the brain of the system.

Types of ML

* Supervised learning
* Unsupervised learning
* Reinforcement learning

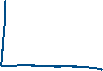
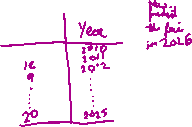
Classification and description of ML:-

* Supervised Learning - Trained on labelled data. Labels are given and they are told to classify on the basis of what labels are there.eg: telling to sort 4-5shapes, email spam detection
* Unsupervised Learning - Classification on the basis of features, trained on unlabelled data. It works by finding out patterns and relationships. It works by creating/generating clusters. Eg: telling to classify the shapes without telling what each shape are.
* Reinforcement - The model learns my making mistakes. It will keep on reiterating the learning by working and iterating till we don’t get the correct answer. It is like a self-teaching system. Eg: a maze solver. It calculates a score (in itself) and then tries to maximize the score.
* Semi-Supervised – The dataset is half labelled. Like all the labels are not present, few labels are absent, the left ones are left out to be determined by the model.

Regression  
The model used to predict the no’s, i.e. how the model predicts values.



Linear regression  
Numerical value prediction in the output.



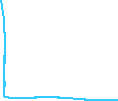
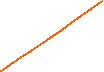
Independent regression🡪 It is not dependent on anything  
Dependent regression 🡪 The price of a commodity in the near future



If there is a difference between high and low-lying values, then we have to normalize the data. We normalize and standardize the data before giving it to the model. It is very imp to detect the outliers as it will be very difficult to deal with them. So, we use boxplot to see the outliers (discussed in last class)



Evaluation Matrix  
We use it to detect how accurately the model is predicting or how good its prediction algorithm is working.   
R^2 method -: Used to see the relation between the target variable and the input. X(independent) and Y(dependent).  
RMSE method -: Root mean square error (it is the most used method for determingin gthe accuracy of the model)   
The gap between the separate pts to get the most accuracy



Random set is important for shuffling.  
If the model’s accuracy is 100% it means it is overpredicting.



MSE = ((1/2m)((i=1 to m)((y1-y1(bar))^2))

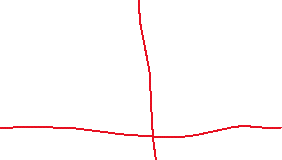
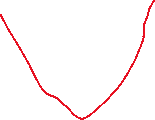
Scalar = (x-(miu))/sigma  
X 🡪independent  
y🡪dependent

Weights and biases

* Weights🡪 It is used to determine the input features having the most effect on the valuation of something.
* Biases🡪 It is the base value of something after removing all the features. There is another definition with overfitting and underfitting part.



Gradient Descent (decreasing the slope of a curve)  
It is the gradual reduction in the error.



=

Error(E) =

Bnew = Bold – Bslope

Bnew = Bold – xslope

For now, we have kept the m value as a constant. We are iterating and seeing for which values, we get the minimum error.

Best fit line 🡺 gradient descent

The loss function had been plotted

We are only changing 1 value at a time so we can get a 2D curve else we will get a 3-D curve which will be difficult to understand.  
We need to ensure that we don’t overshoot and error must be kept minimum  
Random state is a hit and trial method but the gradient descent is a sure shot method. It is a sure method to minimize the error.

So, we need to apply random state and then apply gradient descent.

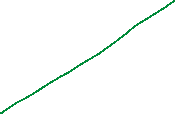
Standard values which are assumed  
learning rate = 0.1 or 0.01  
random state = 42

.ravel() function 🡪 converts 2D array to 1D array(It is an inbuilt library)  
model is predicting in a 2D array

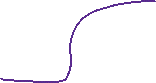
Classification: dividing the data9set of email) label it and then divide it into spam and non-spam. Regression provides continuous numerical values but classification returns proper values(strings)(might be)  
Grouping doesn’t have any labels

Logistic Regression  
It is a statistical method using linear regression model (binary classification-sort of)  
We sort out the data and put them into different labels. We get a lot of probabilities so we need t introduce threshold. We predefine it.

We also introduce a sigmoid function.  
Classification 🡪 Logistic Regression  
Regression 🡪 MSE







Linear, Logistic difference

Label Encoder 🡪 categorical data to numerical data (from last class)

r