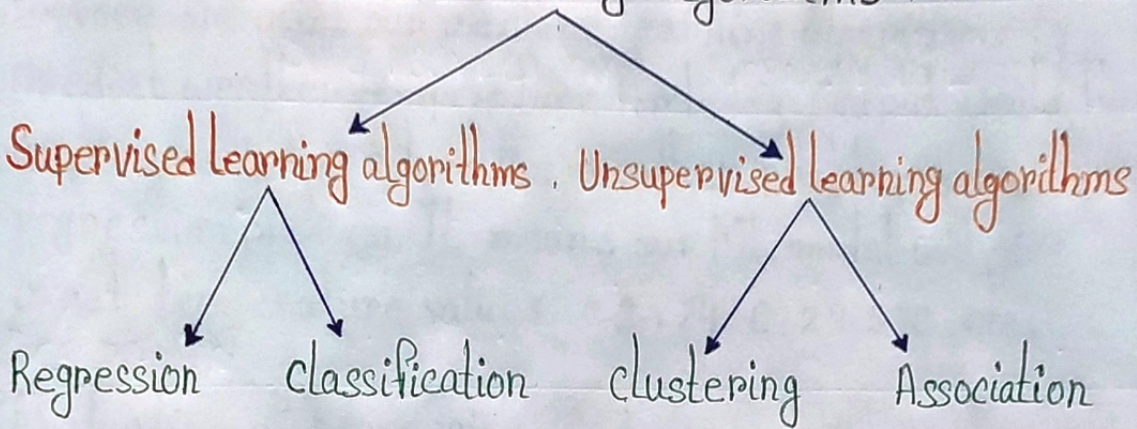
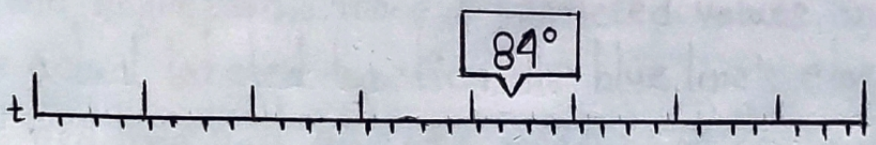
Regression vs. Classification

Machine learning algorithms can be broken down into two distinct types:
Supervised and Unsupervised learning algorithms.



Regression

What will be the temperature tomorrow?



Fahrenheit Meter

Classification

Will it be hot or cold tomorrow?

Fahrenheit Meter

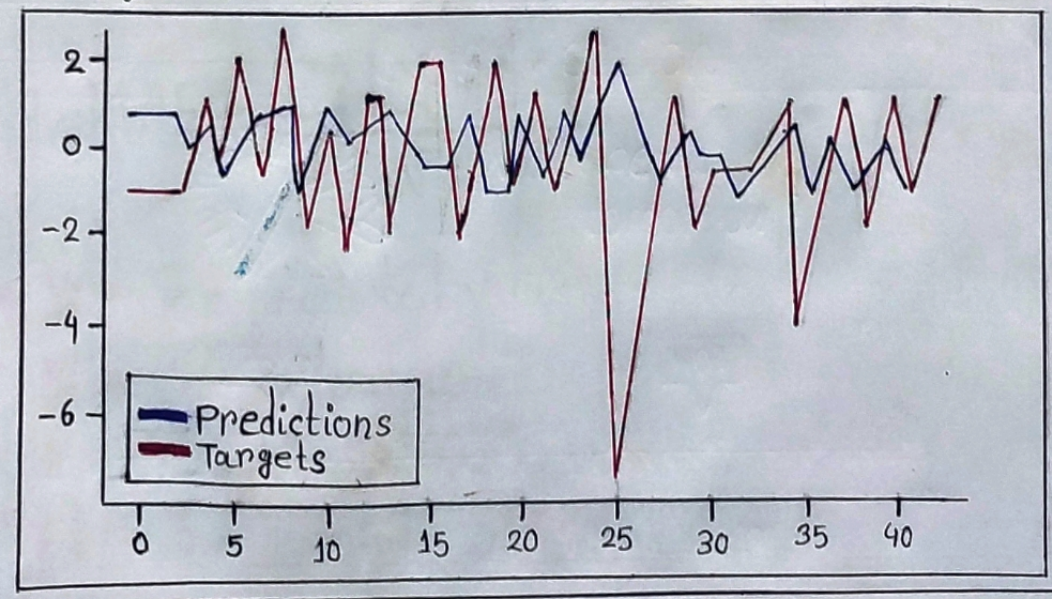
Regression Problems in Machine Learning

Regression is a type of problem that uses machine learning algorithms to learn the continuous mapping function.

Taking the example shown in the last page's image, Suppose we want our machine learning algorithms to predict weather temperature for today. Output would be continuous if we solved the weather problem as a regression problem. It means our ML model will give exact temperature values, e.g., 24°C, 24.5°C, etc.

Evaluation of Regression Models

We estimate the prediction's closeness with the accurately labeled validation/test data to measure the learned mapping function's performance. In the figure below, blue is the regression model's predicted values, and red is the actual labeled function. The blue line's closeness with the red line will give us a measure of How good is our model?



Cost Function for Regression Problem

We define our cost function while building the regression model. It measures the value of the learned values' deviation from the predicted values. Optimizers ensure this error reduces over the progressive iterations, also called epochs. Some of the most common error functions (or cost functions) used for regression problems are:

- Mean Squared Error (MSE)

$$MSE = \frac{\sum_i^N (y_i - y_i')^2}{N}$$

- Root Mean Squared Deviation / Error (RMSD / RMSE)

$$RMSE = \sqrt{\frac{\sum_i^N (y_i - y_i')^2}{N}}$$

- Mean Absolute Error (MAE)

$$MAE = \frac{\sum_i^N |y_i - y_i'|}{N}$$

- y_i = The predicted value.
- y_i' = The actual value.
- N = The total samples.
- $||$ = The absolute value symbol.

Examples of Regression Problems

- Predicting house price based on the size of the house, availability of schools in the area, and other essential factors.
- Predicting sales revenue of a company based on previous sales data of the company.
- Predicting temperature of any day based on wind speed, humidity, and atmospheric pressure.

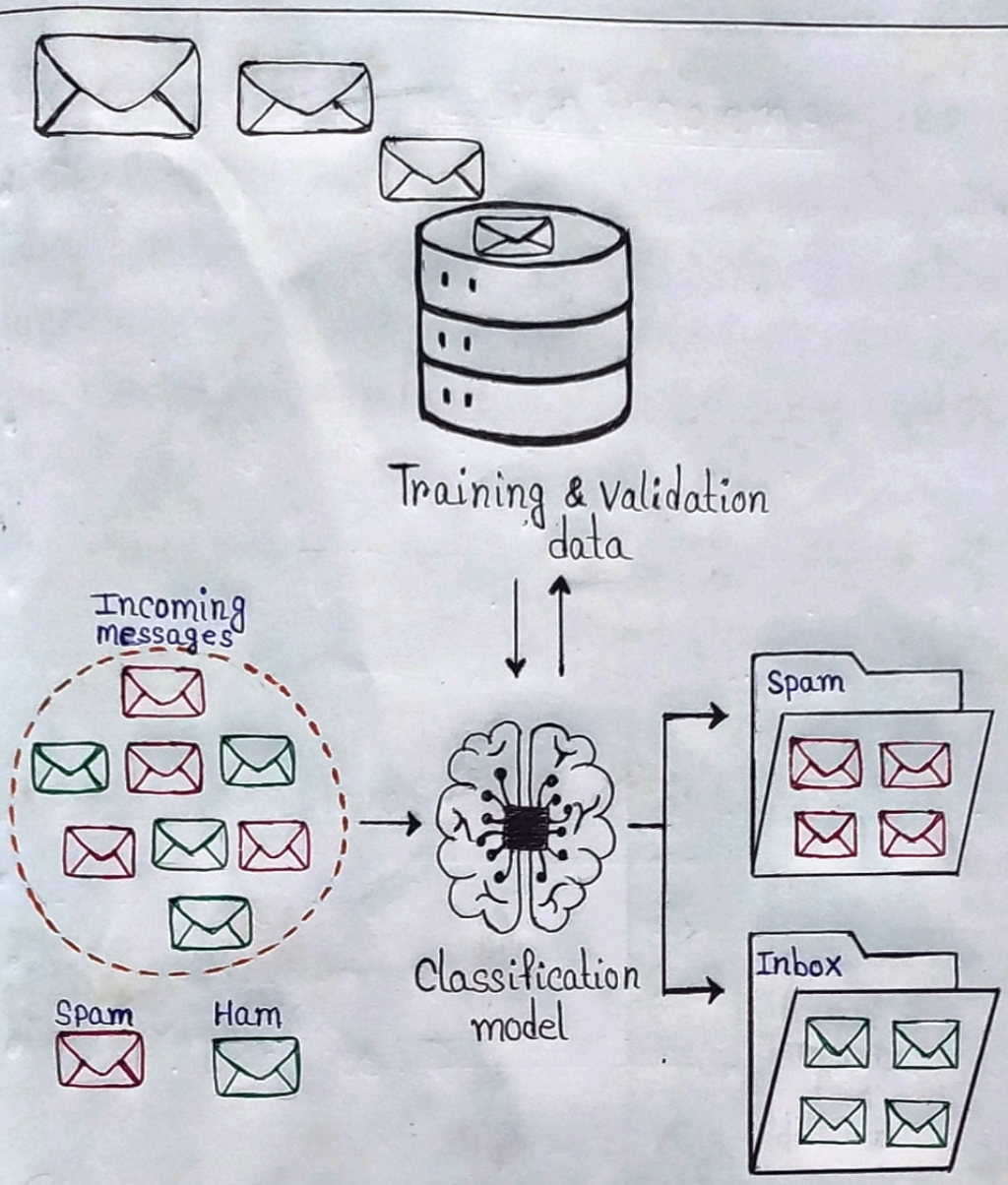
Popular Algorithms Used for Regression Problems

- Linear Regression
- Support Vector Regression
- Regression Tree : Decision Trees, Random Forests, etc.

Classification in Machine Learning

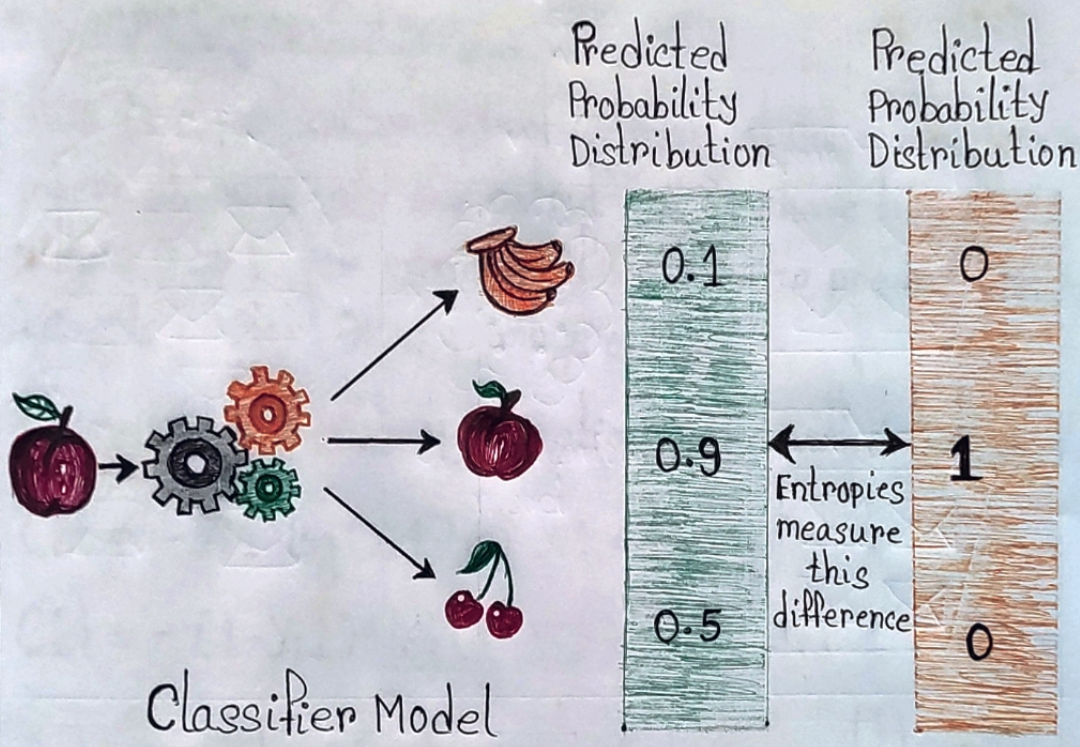
Classification is a supervised machine learning method where the model tries to predict the correct lable of a given input data. In classification, the model is fully trained using the training data, and then it is evaluated on test data before being used to perform prediction on new unseen data.

For instance, an algorithm can learn to predict whether a given email is spam or ham (no spam), as illustrated in the next page.



For example, suppose there are three class labels, [Apple, Banana, cherry]. But the problem is that machines don't have the sense to understand these labels. That's why we need to convert these labels into a machine-readable format. For the above example, we can define Apple=[1,0,0], Banana=[0,1,0], Cherry=[0,0,1]. Once machine learns from these labeled training datasets, it will give probabilities of different classes on the test dataset like this: [P(Apple), P(Banana), P(Cherry)].

These predicted probabilities can be from one type of Probability distribution function (PDF), and the actual (true) labeled dataset can be from another probability distribution function (PDF). If the predicted distribution function follows the exact distribution function, the model is learning accurately.



Cost Function

- **Categorical Cross-Entropy**

Suppose there are M class labels, and the predicted distribution for the ith data sample is :

$$P(Y) = [y_i1', y_i2', \ldots, y_iM']$$

And actual distribution for the sample would be :

$$A(Y) = [y_i1, y_i2, \ldots, y_iM]$$

Cross Entropy (CE_i) =

$$-(y_{i1} * \log(y_{i1}') + y_{i2} * \log(y_{i2}') + \ldots + y_{iM} * \log(y_{iM}'))$$

Categorical Cross Entropy = $-\sum_i^N \frac{CE_i}{N}$

- **Binary Cross-Entropy**

This is a particular case of categorical cross-entropy, where there is only one output that can have two values, either 0 or 1. For example, if we want to predict whether a cat is present in any image or not.

Here, the cross-entropy function varies with the value of y,

$CE_i = -y_{i1} * \log(y_{i1}')$, if $y_{i1} = 1$

$CE_i = -(1 - y_{i1}) * \log(1 - y_{i1}')$, if $y_{i1} = 0$

And similarly, Binary-Cross-Entropy would be averaged over all the datasets.

Most common evaluation metric

- Accuracy
- Confusion Matrix
- F1-Score
- Precision
- Recall etc.

Examples of classification problems

- Classifying if an email is spam or not, based on its content and how others have classified similar types of emails.
- Classifying a dog breed based on its physical features such as height, width, and skin color.
- Classifying whether today's weather is hot or cold.

Algorithms for Classification

- Logistic Regression
- Support Vector Classification
- Decision Tree.

Table

	Regression	Classification
Description	A regression model seeks to predict a continuous quantity.	A classification model seeks to predict some class label
Type of algorithm	Supervised learning algorithm.	supervised learning algorithm.
Type of response variable	Continuous.	Categorial.
How to assess model fit	Root mean squared error	Percentage of correct classifications.