Machine Learning
(Supervised machine learning)

1. <u>Artificial Neural Network (ANN)</u>:

Regression:

a. Simple / Linear regression
b. Multiple linear regression
c. Polynomial regression
d. Support vector regression (SVR)
e. Decision Tree
f. Random Forest

Classification:

a. Logistic regression
b. K- Nearest Neighbour (KNN)
c. Support vector machine (SVM)
d. Decision Tree
    # Entropy, Information Gain and GINI entropy
e. Random Forest

Key concepts:

- **Activation Function**
    a. Sigmoid activation function : Used in o/p layer for Binary classification
    b. Softmax : Used in o/p layer for Multi classification
    c. Relu activation function : Used in hidden layers

- **Loss function / Cost function**
  For Classification:
    a. Binary cross entropy
    b. Categorical cross entropy
    c. Sparse categorical cross entropy
  For Regression:
    a. Mean absolute error (MAE)
    b. Mean square error (MSE)
    c. Root mean square error (RMSE)

- **Underfitting / Overfitting**
- **Bias/ Variance**
- **Hypothesis test : P- value**
- **Regularization :**
    a. L2 regularization and
    b. Dropout regularization
- **Outliers detection technique**
    a. Standard deviation
    b. Box plot
    c. Scatter plot
- **Encoding Technique**
    a. Label encoder
    b. One-Hot encoder
- **Feature scaling**
    a. Standardisation
    b. Normalisation
- **Optimizer Function**
    a. Gradient Descent (SGD) / Stochastic Gradient Descent (SGD) / Mini-Batch SGD

b. SGD with momentum
c. Adaptive Gradient (AdaGrad)
d. RMS propagation / Ada delta
e. Adam optimizer

- **Ensemble method**
  - **a. Bagging**
    - **#** Random Forest
  - **b. Boosting**
    - a. Adaptive Boost (Ada Boost)
    - b. Gradient Boost
    - c. Extreme Gradient Boost (XG Boost)

- **K- fold cross validation**
- **Performance Metrics**

For Classification:

a. Confusion Matrix
b. Recall / Precision / F1 Score

For Regression:

a. R2 , Adjusted R2
b. Mean square error (MSE) , RMSE, Mean absolute error (MAE)

2. <u>Recurrent Neural Network (RNN)</u> : For Time series data analysis

- RNN suitable where data in a particular time series is important.
- **Main Goal :** Predicting the future (Forecasting) and Assign the categories

Algorithm:

a. Long-Sort term memory (LSTM)

Continue ……