# ML ROADMAP

## BEGINNER TO ADVANCED

- **Fundamentals**
  - **Foundational math**

1. Probability
   - Conditional probability
   - Bayes Rule
   - 1D random variable
   - Joint probability distribution

   Resource:
   https://youtu.be/oyyFpAwyt6w?si=IylQrxizLoxKcax7

2. Statistics
   - Statistical Distributions-discrete, continuous,normal
   - Mean,median,mode
   - Sampling distribution
   - Parameter estimation
   - Hypothesis testing
   - Basics of graphs
   - Correlation

   Resource:
   https://youtu.be/M2S_9lyRMvo?si=9ghxDbdBUvVE4C2a
   https://youtu.be/BsV5k-80MJQ?si=K35JF0YZRiU6V7ot

3. Calculus
   - Differentiation
   - Integration

4. Linear Algebra
   - Vectors

- Matrices
- Eigen vectors, Eigenvalues

Resources :

https://youtu.be/uZeDTwWcnuY?si=SBjA5DMAW1W1gW6B

- **Programming skills**
  1. Python
     - Data Types: Dictionary,lists,Sets,Tuples, Strings.
     - Control flow, conditional statements and loops.
     - Function Handling, File Handling and Exception Handling.
     - Libraries for Data Manipulation: Matplotlib, Seaborn, NumPy, SciPy, Pandas.
     - Libraries for Framework Building: Scikit-Learn, Tensorflow(new version includes Keras) and PyTorch.

     Resources-

     1. Python Tutorial For Beginners in Hindi | …
     2. Python basic to advance

  2. DBMS(Optional, For better Database management, For bigger projects)
     - SQL-queries (practice it's questions)

     Resource:

     https://youtu.be/hlGoQC332VM?si=A-ha0oxTV7hf8Fp9

## Core Concepts

- **What is ML?**
  - **Types of ML:**

1. Supervised Learning
2. Unsupervised learning
3. Reinforcement Learning
4. Semi-Supervised Learning
5. Self-Supervised Learning

   Resource-
   https://youtube.com/playlist?list=PLjVLYmrlmjGe-xLyoCdDrt8Nil1Alg_L3&si=CaI0LGzO3_1bysoF

- **Data**

  **- Types of Data**

  - Categorical
  - Numerical

    Resource-

    Understand Data Types with animation | Data Science Cour…

  **- Training, testing and validation data**

    Resource-

    Training Data Vs Test Data Vs Validation Data| Krish N…

  **- Concepts:-**

  1. Overfitting
  2. Underfitting

     Resource-

     a)

     Hindi- Overfitting, Underfitting,Bias And Variance Explained I…

     b)

     Overfitting and Underfitting Explained with Examples in Hind…
  3. Bias and variance
  4. ROC & AUC:
     https://youtu.be/4jRBRDbJemM?si=g5MPyfiUXdhgbDPX
  5. Hyperparameters
  6. PCA

Resources-

a)

Basics Of Principal Component Analysis Part-1 Explained in…

b)

Principal Component Analysis(PCA) Part-2 Explained with S…

c)

Principal Component Analysis Part-3 Explained with Solved …

7. One Hot Encoding
8. Label Encoding

Resources-

a)

One Hot Encoding | Handling Categorical Data | Day 27 | 10…

b)

Encoding Categorical Data | Ordinal Encoding | Label Encod…

9. Evaluation Metrics

Resources-

a)

Confusion Matrix ll Accuracy,Error Rate,Precision,Recall Ex…

b) Machine Learning Fundamentals: Sensitivity and Specificity

c) The Sensitivity, Specificity, Precision, Recall Sing-a-Long!!!

d) ROC and AUC, Clearly Explained!

10. Handling Missing Values

Resource-

4.3. Handling Missing Values in Machine Learning | Imputati…

**- Data Manipulation(Hands-on workshops would be conducted)**

- Data cleaning
- Data transformation
- Feature Scaling
- Feature engineering
- Data exploration

Resource : https://youtu.be/4hYOkHijtNw?si=pPN8rgxDKM3cDsbI

- **MODELS**
    - How does the model work?

1. ALGORITHMS
    - Logistic regression
        - Machine Learning Tutorial Python - 8: Logistic Regression (Binary …
    - Linear regression
        - Simple Linear Regression | Code + Intuition | Simplest Explanation i…
        - Car Price Predictor Project | Machine Learning | Linear Regression
    - SVM
        - https://www.geeksforgeeks.org/support-vector-machine-algorithm/
        - Support Vector Machines Part 1 (of 3): Main Ideas!!!
        - Lecture 2.6 | Support Vector Machine (SVM) | Complete Explanatio…
    - KNN
        - https://www.geeksforgeeks.org/k-nearest-neighbours/
        - Lec-7: kNN Classification with Real Life Example | Movie Imdb Exa…
        - K Nearest Neighbors Part 2 - Working with a Real world dataset
        - StatQuest: K-nearest neighbors, Clearly Explained
    - Decision Trees
        - Decision Tree Classification in Machine Learning | Decision Tree in …
        - Decision and Classification Trees, Clearly Explained!!!
    - Random Forests
        - https://youtube.com/playlist?list=PLKnIA16_RmvZyqP3WGUo7iVzilIea_1bp&si=ja7P8NZoRWxKFwY0
        - StatQuest: Random Forests Part 1 - Building, Using and Evaluating
        - Machine Learning Project | Employee Attrition Prediction | Random …
    - Naive Bayes
        - Naive Bayes, Clearly Explained!!!
        - https://www.geeksforgeeks.org/naive-bayes-classifiers/

## DEEP LEARNING

- Introduction
- Neuron
- Neural networks
- ANN, CNN, RNN(GRU, LSTM), GAN
- Pytorch, Tensorflow 2.0(including Keras)
- Gradient descent
- Dropout regularization

  Resource:
  https://youtube.com/playlist?list=PLKnIA16_RmvYuZauWaPlRTC54KxSNLtNn&si=qxKYDt6qy9wL-rj3

## DEPLOYMENT

- Google cloud platform
- Flask

  Resource-
  https://youtube.com/playlist?list=PL-osiE80TeTs4UjLw5MM6OjgkjFeUxCYH&si=QYubpMowuZRk1IOs
- Django
- Streamlit-

  Resource-

  1) Complete Streamlit Python tutorial in One Video I Machine Learni…
  2) Machine Learning Model Deployment Using Streamlit

# NOTE

1. Practice self made mini-projects from GeeksforGeeks, Kaggle and other platforms.
2. Hands-on sessions would be conducted.
3. Any new topic which you come across and is not mentioned in the roadmap, just use ChatGPT and Google to know more about the topic.
4. Any topic you feel is important and is not mentioned, you can tag the Circle Managers on the group to add it.
5. For a certain topic choose one suitable resource out of the given choices and stick with it.