**Topics:**

1. Pandas, Numpy and Matplotlib

<https://www.dataquest.io/m/289-introduction-to-numpy/> <https://cloudxlab.com/blog/numpy-pandas-introduction/>

<https://www.hackerearth.com/practice/machine-learning/data-manipulation-visualisation-r-python/tutorial-data-manipulation-numpy-pandas-python/tutorial/>

<https://towardsdatascience.com/data-visualization-using-matplotlib-16f1aae5ce70>

1. Introduction to Machine Learning
   1. Need for Machine Learning,
   2. Basic Principles
   3. Applications
   4. Challenges
   5. Types of Machine Learning:
      1. Supervised Learning
      2. Unsupervised Learning
      3. Reinforcement Learning
2. Supervised Learning
   1. Naive Bayes algorithm: Mathematics behind naive bayes algorithm. Numerical Example of naive bayes algorithm. Practical implementation of naive bayes algorithm
   2. Linear Regression
      1. Ordinary Least Square method
      2. Simple Linear Regression
         1. Making predictions
         2. Cost function(SSE)
         3. Gradient descent
         4. Training
         5. Model evaluation
      3. Multivariable Linear Regression
      4. Polynomial Regression
      5. Applications

Refer following free course on datacamp. Regression models: fitting them and evaluating their performance

* 1. Logistic Regression
     1. Comparison to linear regression
     2. Types of logistic regression
     3. Binary vs Multi class logistic regression
     4. Sigmoid Function
     5. Decision boundary
     6. Probability to Class labels mapping
     7. Log Loss function
     8. Gradient descent
     9. Training and Testing
     10. Multiclass logistic regression
         1. Softmax activation
     11. Applications

<https://medium.com/@anishsingh20/logistic-regression-in-python-423c8d32838b>

<https://www.datacamp.com/community/tutorials/understanding-logistic-regression-python>

<https://towardsdatascience.com/logistic-regression-python-7c451928efee>

<https://towardsdatascience.com/building-a-logistic-regression-in-python-step-by-step-becd4d56c9c8>

<https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html>

* 1. K Nearest Neighbors
     1. Regression and Classification
     2. Deciding values of K
     3. Training and Testing
     4. Applications
  2. Training Testing Splits and Cross-Validation
  3. <https://www.analyticsvidhya.com/blog/2018/03/introduction-k-neighbours-algorithm-clustering/>

1. Goodness of Model
   1. Confusion Matrix
   2. Accuracy
   3. Precision
   4. Recall
   5. F1 core
   6. Sum of Squared Error
   7. Root Mean Squared Error
   8. R2 Score

<https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/>

**Sample Questions**

1. Sigmoid is given as h = 1/1+e^-z. Now x = [5.72,6.1,4.2] and y = [in,in,out]. What would be the h. Assume all values of weight vectors are zero.
2. Which of the following is an application of supervised learning?
   1. Around 500000 public opinions about political elections are given in a small text format and estimate which party will win the election.
   2. Estimate price of a Second-hand car based on given conditions.
   3. Divide customers based on their buying habits
   4. Estimate which mutual fund is buyable.
3. What is Type I error? Explain in 1 line.
4. Write a python function named ED for finding Nearest neighbors using Euclidean Distance.