Introduction to Data Science Course Curriculum

1. **Introduction to data science**
   a. Overview
   b. Terminology in data science
   c. Application of data science

2. **Basics Python for data science**
   a. Introduction
   b. Operators
   c. Variables
   d. Variables and data types
   e. Conditional statements
   f. Looping
   g. Function
   h. Data structure
   i. Lists, Dictionary and Tuple
   j. Regular Expression
   k. File Reading (CSV, Excel etc.)
   l. Basics Class and Objects concepts

3. **Numpy**
   a. Introduction
   b. Numpy.zeros() and numpy.ones() in python
   c. Numpy.reshape() in python
   d. Numpy.hstack() and numpy.vstack()
   e. Numpy.asarray()
   f. Numpy.arange()
   g. Numpy.linspace() and numpy.logspace()
   h. Indexing and Slicing
   i. Numpu statistical function
   j. Numpy.dot(): dot product in python
   k. Numpy matrix manipulation with np.matmul()

4. **Pandas**
   a. Introduction
   b. Data structure
   c. Data types

8. **Linear Algebra**
   a. Introduction
   b. Scalars, Vectors, Metrics and Tensors
   c. Multiplication of Metrics and vectors
   d. Identity, inverse etc.
   e. Eigenvalues and Eigenvectors

9. **Calculus**
   a. *https://www.youtube.com/playlist?list=PLZHQObOWTQDMsr9K-rj53DwVRMYO3t5Yr&fbclid=IwAR2tbiSGEvuPqI5RMWvINEFIj88DcobnFX58KWacjPLxTo2VNgTpHHtrUkc*

10. **Predictive Modeling**
   a. Introduction to predictive modelling
   b. Types of predictive modelling
   c. Hypothesis Generation
   d. Data Extraction
   e. Data Exploration
      i. Variables
      ii. Data Types
      iii. Univariate Analysis
         1. Descriptive Analysis
         2. Discreate variable barplot, Countplot, Boxplot, Distplot
      iv. Bi-Variate Analysis
         1. Scatter plot
         2. Correlation Plot
         3. Heat Map
      v. Feature Cleaning
         1. Missing Value imputation
         2. Listwise deletion
         3. Mean/Median/Mode imputation
         4. End of distribution

    c. Crossing

    d. Ratio proportion

    e. Cross product

    f. Feature learning by rate

    g.

   viii. Feature Selection

    1. Filter method

     a. Variance

     b. Correlation

     c. Chi-square

     d. Information Value

    2. Wrapper method

     a. Forward Selection

     b. Backward elimination

     c. Genetic algorithm

     d. Exhaustive feature selection

    3. Embedded Method

     a. Lasso(L1)

    4. Hybrid Method

     a. Recursive Feature Selection

     b. Recursive Feature addition

11. Machine Learning Important Algorithms and related concepts

 a. Linear Regression

 b. Logistics Regression

 c. Decision Tree

 d. SVM

 e. Naïve Bayes

 f. KNN

 g. K-means

 h. Random Forest

 i. Dimensionality Reduction

 j. Gradient Boosting algorithms

 k. Clustering algorithms

******************Some More important Concepts******************

- Introduction to evaluation metrics
- Confusion Matrix
- Accuracy
- Precision, Recall and F-score
- Thresholding
- AUC and ROC
- Log Loss
- Evaluation matrix for regression
- Means squared
- RMSE