**Python Machine Learning in 7 Days**

**Section 1: Enter the Machine Learning World!**

This section will cover introduction to machine learning, why is this course good for them and how its structured, setting up the environment and introduction to libraries and a short introduction to types of machine learning models.  
  
The student will then work on a pre-built Machine Learning model and train them to build more efficient models.

* The Course Overview
* Setting Up Your Machine Learning Environment
* Exploring Types of Machine Learning
* Using Scikit-learn for Machine Learning
* Assignment – Train Your First Pre-built Machine Learning Model

**1.1 The Course Overview**

This video provides an overview of the entire course.

**1.2 Setting Up Your Machine Learning Environment**

This video teaches you how to set up your machine learning environment.

* + Download and Install Anaconda
  + Get introduced to Jupyter Notebook

**1.3 Exploring Types of Machine Learning**

In this video, we will talk about different types of Machine Learning Models.

* + Learn about Supervised Learning
  + Understand what is Unsupervised Learning
  + Understand what is Reinforcement learning

**1.4 Using Scikit-learn for Machine Learning**

This video introduces Scikit-learn library.

* + Learn about scikit-learn
  + Look at the Data Representation in Scikit-Learn
  + Learn about the Estimator API

**1.5 Assignment – Train Your First Pre-built Machine Learning Model**

This video introduces the first assignment.

* + Look at what is to be done to solve the first assignment

**Section 2: Build Your First Predicting Model**

This section starts from where the last module left off and goes deeper into supervised learning. Introduces the basic design of a machine learning model and covers topics like cost function, gradient descent, dependent and independent variables and then proceeds into explaining these with the help of linear regression. We will then see how to implement a linear regression model in python and predict height. The student will then use the acquired knowledge to make predictions of a power plant’s energy output.

* Supervised Learning Algorithm
* Architecture of a Machine Learning System
* Machine Learning Model and Its Components
* Linear Regression
* Predicting Weight Using Linear Regression
* Assignment – Predicting Energy Output of a Power Plant

**2.1 Supervised Learning Algorithm**

In this video, we will understand what Supervised learning is.

* + Look at common Examples of Supervised learning
  + Learn Regression as a type of Supervised learning algorithm
  + Understand Regression as a type of Classification learning algorithm

**2.2 Architecture of a Machine Learning System**

In this video, we will see the general architecture of machine learning algorithms.

* + Understand about Hypothesis
  + Understand about Parameters
  + Learn how to use a Cost Function

**2.3 Machine Learning Model and Its Components**

In this video, we will learn about the important components of a model.

* + Learn how Error Minimization works
  + Learn how Gradient Descent works
  + Intuition for Gradient Descent

**2.4 Linear Regression**

# This video explains Linear Regression in details.

* + Look at steps for Linear Regression training
  + Understand about Batch Gradient Descent

**2.5 Predicting Weight Using Linear Regression**

In this video, we will learn how to use Scikit-learn for Linear Regression.

* + Basic operations and syntax of Scikit-learn
  + Use Jupyter Notebook
  + Create Linear Regression Model

**2.6 Assignment – Predicting Energy Output of a Power Plant**

We will look at how to solve the assignment.

* + Step by step guide on solving the assignment

**Section 3: Image Classification Using Supervised Learning**

We start with reviewing Tuesday’s assignment by looking at an ideal solution and results. This section then goes into logistic regression, another form of supervised learning, and also introduces a very common machine learning problem of classification. It then moves into Support vector machines which performs really well when compared to other classification techniques. We then cover different types of kernels in SVM.  
  
The student is then supposed to use his implementation knowledge from logistic regression implementation and output a better result using (Support Vector Machines) SVM.

* Review of Predicting Energy Output of a Power Plant
* Logistic Regression
* Classifying Images Using Logistic Regression
* Support Vector Machines
* Kernels in a SVM
* Classifying Images Using Support Vector Machines
* Assignment – Start Image Classifying Using Support Vector Machines

**3.1 Review of Predicting Energy Output of a Power Plant**

In this video, we will learn how to solve Section 2 assignment

* + Download and manipulate the data
  + Fit a linear\_model. LinearRegression() model
  + Find out the mean squared error and variance score and plot regression line on the scatter plot

**3.2 Logistic Regression**

This video will teach you about Logistic Regression

* + Learn about Classification problem
  + Learn how Hypothesis function in Logistic Regression looks
  + Understand Cost Function and Gradient Descent in Logistic Regression

**3.3 Classifying Images Using Logistic Regression**

This video aims at creating a logistic regression model to classify handwritten digits

* + Understand the dataset
  + Import the dataset into Jupyter Notebook and convert into a numpy array
  + Use logistic regression model

**3.4 Support Vector Machines**

In this video, we will learn about Support vector machines

* + Learn what a support vector machine is
  + Learn how Support Vector Machines work

**3.5 Kernels in a SVM**

In this video, we will look at the Kernels in SVM

* + Understand the need for Kernels in SVM
  + We will look at nonlinear problems
  + Understand how Kernels in SVM solve this problem

**3.6 Classifying Images Using Support Vector Machines**

In this video, we will see how SVM works for Classifying images

* + Understand the dataset
  + Import the dataset into Jupyter Notebook and convert into a numpy array
  + Use Support Vector Machines

**3.7 Assignment – Start Image Classifying Using Support Vector Machines**

In this video, we will look at how to solve the assignment

* + Step by step guide on solving the assignment

**Section 4: Improving Model Accuracy**

We start with reviewing previous day’s assignment by looking at an ideal solution and results. This module takes a pause from introducing any new model and focuses on improving the models and re-focuses on the broader picture of applying a machine learning technique. This module will talk about precision and recall, accuracy, and training vs test vs validation set results.  
  
The students are then encouraged to improve their results from last two assignments and measure their own models and improve them with different iterations.

* Review of Classifying Images Using Support Vector Machines
* Model Evaluation
* Better Measures than Accuracy
* Understanding the Results
* Improving the Models
* Assignment – Getting Better Test Sample Results by Measuring Model Performance

**4.1 Review of Classifying Images Using Support Vector Machines**

In this video, we will learn how to solve Section 3 assignment

* + Download and manipulate the data
  + Open, resize and get the features of each image using the PIL library
  + Use SVM to train the model and find out model accuracy

**4.2 Model Evaluation**

This video aims at explaining the concept behind Model evaluation

* + Understand Parameters and Hyperparameters
  + Learn test and validation datasets
  + Understand K-fold Cross Validation as a way to support model evaluation

**4.3 Better Measures than Accuracy**

In this video, we will learn about different ways of measuring model performance

* + Learn the details of Classification Metrics
  + Explore Regression Metrics

**4.4 Understanding the Results**

This video will learn how to make sense of model evaluation

* + Understand what Bias and underfitting means
  + Understand what Variance and overfitting mean
  + Learn why we strive for a Bias Variance Tradeoff

**4.5 Improving the Models**

We will look at ways to improve the model

* + Understand how we can use Regularization
  + Learn feature scaling as another way to improve the model

**4.6 Assignment – Getting Better Test Sample Results by Measuring Model Performance**

We will look at how to solve the assignment

* + Step by step guide on solving the assignment

**Section 5: Finding Patterns and Structures in Unlabeled Data**

We start with reviewing previous day’s assignment by looking at an ideal solution and results. Module will then introduce unsupervised learning and dwells into K means for better understanding. Students will learn how to measure K means algorithms performance and how to choose an ideal solution. The students will then use this expertise to write their own clustering implementation to cluster plant samples together.

* Review of Getting Better Test Sample Results by Measuring Model Performance
* Unsupervised Learning
* Clustering
* K-means Clustering
* Determining the Number of Clusters
* Assignment – Write Your Own Clustering Implementation for Customer Segmentation

**5.1 Review of Getting Better Test Sample Results by Measuring Model Performance**

In this video, we will learn how to solve Section 4 assignment.

* + Download and manipulate the data
  + Check if the model is overfitting by using cross validation
  + Use regularization to reduce overfitting

**5.2 Unsupervised Learning**

In this video, We will take a closer look at unsupervised learning.

* + In this video, we will revisit unsupervised learning
  + We will also look at some examples of application of unsupervised learning

**5.3 Clustering**

In this video, we will look at a clustering as a unsupervised technique.

* + In this video we will see how clustering works
  + We will look at a popular clustering algorithm called k-means clustering

**5.4 K-means Clustering**

We will then look at K means clustering one of the most popular clustering techniques.

* + In this video we will look at the details of how K-means work including the cost function
  + We will also be looking at the optimization objective
  + We will explore ways to improve the performance of k-means clustering by using different initialization techniques

**5.5 Determining the Number of Clusters**

We will look at ways to optimize our k means algorithm by finding the number of nodes.

* + In this video we will look at ways to find out the optimal number of clusters to be used
  + Two popular ways to find out the optimal number of clusters is elbow method and business rules

**5.6 Assignment – Write Your Own Clustering Implementation for Customer Segmentation**

We will end the section with an assignment on Customer Segmentation.

* + Look at how to use pandas library
  + Understand the dataset to be used for this assignment
  + Learn the steps to solve the assignment

**Section 6: Sentiment Analysis Using Neural Networks**

We start with reviewing previous day’s assignment by looking at an ideal solution and results. Module will then introduce unsupervised learning and dwells into K means for better understanding. Students will learn how to measure K means algorithms performance and how to choose an ideal solution. The students will then use this expertise to write their own clustering implementation to cluster plant samples together.

* Review of Clustering Customers Together
* Why Neural Network
* Parts of a Neural Network
* Working of a Neural Network
* Improving the Network
* Assignment – Build a Sentiment Analyzer Based on Social Network Using ANN

**6.1 Review of Clustering Customers Together**

In this video, we will learn how to solve Section 5 assignment

* + Download and manipulate the data
  + Find out the optimal number of k-means clusters using the elbow method (using scipy.spatial.distance.cdist for distance)
  + Perform k-means++ with the optimal number of clusters and report values of the cluster centers

**6.2 Why Neural Network**

We will be looking at Neural Network from a supervised learning perspective.

* + Understand the advantages of using a neural network and try to understand its popularity
  + Look at the applications of neural network

**6.3 Parts of a Neural Network**

In this video, we will be looking at the parts of a neural network in more details.

* + Learn the components of a neural network
  + Learn about activation functions
  + Understand how a neural network produces its output

**6.4 Working of a Neural Network**

In this video, we will be looking at the working of a neural network.

* + look at the cost function of a neural network
  + Revisit gradient descent
  + Learn the algorithm called backpropagation

**6.5 Improving the Network**

In this video, we are going to look at ways to improve a neural network model.

* + Understand how to change the activation function
  + Find ways to prevent overfitting
  + Look at early stopping as a way to prevent overfitting

**6.6 Assignment – Build a Sentiment Analyzer Based on Social Network Using ANN**

We will end the section with an assignment on Sentiment Analysis.

* + Look at a text mining technique called as the bag of words feature
  + Look at words which do not support my text classification task at hand called known as stop words
  + Understand the dataset and look at the steps to solve the assignment

**Section 7: Mastering Kaggle Titanic Competition Using Random Forest**

The module will cover Random Forest and thus cover decision trees and will encourage the student to make a Kaggle submission and introduce them to the online machine learning challenges and community

* Review of Building a Sentiment Analyser ANN
* Decision Trees
* Working of a Decision Tree
* Techniques to Further Improve a Model
* Random Forest as an Improved Machine Learning Approach
* Weekend Task – Solving Titanic Problem Using Random Forest

**7.1 Review of Building a Sentiment Analyser ANN**

In this video, we will learn how to solve Section 6 assignment

* + Download and manipulate the data
  + Remove the Stop words from the data and create a bag of words feature matrix using CountVectorizer()
  + Train a Neural Network and print out the confusion matrix

**7.2 Decision Trees**

In this video, we are going to look at a machine learning algorithm called Decision trees

* + Understand decision trees better with an example
  + Explore some advantages of using decision trees

**7.3 Working of a Decision Tree**

In this video, we will be looking at the Working of a decision tree

* + Find out how a decision tree is split
  + Learn about the criteria to split the nodes

**7.4 Techniques to Further Improve a Model**

In this video, we are going to look at ways to avoid overfitting in a decision tree

* + Understand the problem of overfitting in Decision Trees
  + Look at ways to avoid overfitting in Decision Trees

**7.5 Random Forest as an Improved Machine Learning Approach**

In this video, we are going to look at ways to improve model performance by using Random Forest

* + Understand the technique called ensemble learning which improves the performance our machine learning algorithm
  + Look at a powerful machine learning technique called random forests

**7.6 Weekend Task – Solving Titanic Problem Using Random Forest**

In this video, we are going to look at the dataset and what the weekend task looks like

* + Understand the dataset we are going to use
  + Predict the chance of survival which is dependent variable
  + Complete the assignment we Download the test and train datasets from the Kaggle Site in the Kaggle Titanic competition section