SUMMER ANALYTICS

MODULE 04: INTRODUCTION TO MACHINE LEARNING

Hello everyone!

Kudos to you on completing the first three modules!

In this module we will begin our journey to understand the basic concepts of Machine Learning and also you'll implement some models yourself. Also, optional content is added in case you need better clarity on some topics.

TOPICS:

1. **Machine Learning and its types**:
   Understanding the intuition behind what Machine Learning and its types will be very handy when we try to solve a problem ourselves.

TASK 1	What is Machine Learning	A short video to build an intuition what ML is
TASK 2	ML Basics	Understanding ML with a use-case
TASK 3	Types of ML	We should be aware of what category of ML will our problem fall into
TASK 4	Classification VS Regression	The burning hot question of Supervised Learning gets answered using this video

2. **Linear Regression**
   Let's build our understanding of the first type of ML model!

TASK 1	Basic concepts in Linear Regression (2.1 - 2.7)	Basic concepts include Gradient descent and Cost Function for a Linear Regression model
TASK 2	Linear Algebra theory (3.1 to 3.6)	Linear Algebra theory required for implementation
TASK 3	Linear Regression article	A summary of the process for review
TASK 4	Linear Regression with multiple variables (4.1 to 4.7)	A more generalized implementation of Linear Regression

TASK 5	Scikit-learn	A blog on the powerful sklearn library
TASK 6	Install Scikit-learn	Anaconda comes with sklearn pre-installed. You could check the latest version at the given link and upgrade it too. (optional - for users without Anaconda)
TASK 7	Train Test Split	A blog on applying the famous train_test_split() in python
TASK 8	Linear Regression in sklearn	Implement your first model Linear Regression model*

Note : At 0:40 in Task 7, use sklearn.model_selection instead of sklearn.cross_validation.
To know about sklearn.datasets.load_boston(), refer to this link.

3. Logistic Regression
Now, we will tackle a Classification based problem using Logistic Regression.

TASK 1	Basic Concepts in Logistic Regression (6.1 - 6.7)	Basic concepts in Logistic Regression
TASK 2	Logistic Regression with Scikit-Learn (binary classification)	A walk-through video for Binary classification
TASK 3	Logistic Regression with Scikit-Learn (multiclass classification)	A walk-through video for Multiclass classification

4. K-Nearest Neighbours
KNN is one of the easiest to grasp yet a very powerful algorithm in Machine Learning.

TASK 1	How KNN works	Intuition for KNN
TASK 2	Feature Scaling (Standardization vs Normalization)	ML models can give pretty bad results if not scaled properly. This blog is on Feature Scaling
TASK 3	KNN review	A summary of the application of KNN in sklearn for review

5. Some other concepts

TASK 1	Overfitting and Underfitting video	Understanding the problem of overfitting and underfitting
TASK 2	Overfitting and Underfitting read	An article to get a better understanding
TASK 3	Bias-Variance tradeoff	Introduction to the Bias-Variance tradeoff
TASK 4	Bias-Variance tradeoff graphical	Interesting graphics for grasping the trade-off
TASK 5	Regularization (7.1 - 7.4)	Regularization is a way of solving overfitting. Let's see how!
TASK 6	Ridge regression	An introduction to ridge regression
TASK 7	Lasso regression	An introduction to lasso regression

6. **Optional content**

TASK 1	[The intuition of linear regression](#)	
TASK 2	[Generalization, overfitting, and underfitting](#)	
TASK 3	[Regression metrics](#)	
TASK 4	[KNN using sklearn](#)	

Congratulations, you've completed the content of this module!
Time for some self-check Assignments

7. **Self-Check Assignments**

TASK 1	[Linear Regression Self-Check Notebook](#) [Dataset link](#)	Try completing the EDA and model building and refer solutions if you are stuck
TASK 2	[Linear Regression Self-Check Notebook: Solutions](#)	Solutions of the Self-Check Assignment
TASK 3	[Logistic Regression Self-Check Notebook](#) [Dataset link](#)	Try completing the EDA and model building and refer solutions if you are stuck
TASK 4	[Logistic Regression Self-Check Notebook: Solutions](#)	Solutions of the Self-Check Assignment
TASK 5	[KNN walk-through](#) [Dataset link](#)	This notebook contains the code already written. Run it to see how KNN is modeled
TASK 6	[KNN Regression Self-Check Notebook](#) [Dataset link](#)	Try completing the EDA and model building and refer solutions if you are stuck
TASK 7	[KNN Regression Self-Check Notebook: Solutions](#)	Solutions of the Self-Check Assignment

8. This component is graded. Feel free to use the internet to understand the concepts better. Do these questions independently for the best results.

Click [here](#) to attempt the test.