TRAINING CURRICULUM

COURSE TITLE	Data Science/Machine learning curriculum
COURSE DURATION	TBD
SKILL LEVEL	Beginner - Intermediate
AUDIENCE	Data scientist and machine learning enthusiast

Course Overview:

Data Science and Machine learning are great fields and everyone seems to be talking about it, but what exactly is Data Science? What is machine learning? how does it connect back to other technology fields and what is it impacts at large in building systems to scale and getting maximum business value? - There are also questions like where is this Data coming from? How do we make the best use of them and what are the systems and technologies used?

These and more are some of the questions we hope to answer in this course.

Learning Outcomes:

Participants will;

- understand the roles and responsibility of a Data Scientist.

- Deep dive knowledge of Data Visualization.

- Technologies used in Data Science and machine learning.

Knowledge Requirements:

Basic knowledge of Data science.

Hands on experience with languages like Python, R.

Course Outline:

Introduction to Big Data	
1.1	Big Data, ETL pipelines, Data Mining, Predictive Analytics
2.1	Fundamentals of Data Mining

2.2	Dataset types, Data preprocessing, Similarity, Data exploration

Introduction to R/Python programming And Regression

3.1	R/Python basics
3.2	R/python data types
3.3	R/python language features
3.4	R/python visualization
4.1	Cost Functions
4.2	Logit Function
4.3	Decision Boundaries

Introduction to AWS Machine learning platforms and Data Exploration and Visualization

5.1	AWS, S3, AWS ML Data sources
6.1	Exploration
6.2	Visualization
6.3	Segmentation, R, Python

Introduction to predictive Analytics and classification / Evaluation of predictive Models / Ensemble Methods

7.1	Predictive Analytics, Classification, Decision Trees, Gini index
7.2	Entropy, Training/Test splits
8.1	Accuracy, Precision, Recall, ROC, AUC, Cross-validation, Bias/Variance Tradeoff
9.1	Binomial Distribution, Bagging, Boosting, Random forests, Adaboost.

Deployment Machine Learning Models / Parameter Tuning

10.1	Real-time Prediction, Batch prediction
10.2	REST endpoints, Security
11.1	Model Tuning, Parameters, Bias/Variance Tradeoff

CONCLUSION

Learning Aids:

- Videos
- Slide Presentations
- Case Studies
- Group Assignments/Activities.