

Part 2: Training Programme Details (ADS)

Section A: Course Details

1	Course Title	Applied Data Science
2	Type	Technical
3	Training Methodology	Classroom (physical; with virtual as option)
4	Skill Area	Data Analytics & Data Science Lifecycle Data Analysis Techniques Develop Predictive Models with various Approaches Build Recommendation Engines Analytics Presentation Skills
5	Duration	4.0 days / 28 hours
6	Certification	Certificate of Completion
7	Certification Body	N/A
8	Course Overview	<p>Descriptive and diagnostic analytics no longer sufficient. More and more companies are moving into predictive and prescriptive analytics to position themselves in a more competitive edge. Data Science Professions are the brains behind this.</p> <p>This course enables participant to develop a deep and profound understanding of Big Data Analytics (BDA) and the Data Science lifecycle, and be able to develop predictive models and recommendation engines.</p>
9	Prerequisites	Diploma / Degree or Equivalent
10	Course Objective	Upon completion of this course, participants will be able to develop a deep and profound understanding of Big Data Analytics (BDA) and the Data Science lifecycle, and be able to develop predictive models and recommendation engines.
11	Learning Outcome	By the end of the training, participants will be able to develop a deep and profound understanding of Big Data Analytics (BDA) and the Data Science lifecycle, and be able to develop predictive models and recommendation engines.
12	Course Content	Day 1: Introduction to Data Science <ul style="list-style-type: none">• What is Data Science• Analytics Spectrums• Important of Data

- Datasets and data types
- Knowledge Discovery from Data
- Data Science Lifecycle
- Commonly used data sciences techniques

Introduction Machine Learning and MLLib

- What is Machine Learning?
- Top 5 MLLib

Data Acquisition and Cleansing

- Data Acquisition and preparation

Exploratory Data Analysis

- What is exploratory data analysis?
- The basic EDA methods

Day 2:

Types of Machine Learning and Regression models

- Type of Machine Learning
- Regression models

Regression models

- Regression models
- Performance Evaluation for Regression models
- Regression: When and Which

Classification models

- Classification models
- Performance Evaluation
- Classification: When and Which

Day 3:

Clustering models

- Clustering Models
- Classification: When and Which

Geo Clustering

- What is Geo Clustering
- Geo Spatial Data
- Geo Clustering

Graph Analysis

- What is Graph Analysis
- Graph Analysis
- Graphs in Graph Analysis
- Graph Analysis Demo

Recommender System

- Introduction to Recommender System
- Recommender System Benefits Service Providers
- Types of Recommender System
- Types of Collaborative Filtering
- Other Methods of Recommender System
- Recommender System Using Apriori Algorithm

Day 4:

Natural Language Processing (NLP)

- What is Natural Language Processing

		<ul style="list-style-type: none"> • Natural Language Processing Overview • Text Processing and NLP • Text Classification • Text Matching • Important Libraries for NLP (using Python) <p>Sentiment Analysis</p> <ul style="list-style-type: none"> • What is Sentiment Analysis? • Sentiment Analysis <p>Model Evaluation Metrics</p> <ul style="list-style-type: none"> • What is Model Evaluation Metrics • Types of Model Evaluation Metrics • Test Sets • Learning Curve • Limitations of Single Training/Test Partition • Stratified Sampling • Cross Validation • Confusion Matrix • Other Accuracy Metrics • Precision/Recall Curves • Text Classification – Classifier Evaluation
13	Learning Activities	Lecture, Practical Exercise, Case Studies, Learning Activities, Video Presentation, Training
14	Target Group	This course caters to those with experience in Data Science and have intentions of becoming full-fledged Data Scientists. Participants should preferably have some knowledge in Python or R and are recommended to have completed the Applied Data Analytics course.

Detailed Daily Training Schedule (Course Content / Hours)

No.	Content/Activity	Objectives	Outcome	Hours
1	Day1 [9am-11am] - Introduction to Data Science	This section covers Analytics Spectrums, Data Science Lifecycle, Commonly used data sciences techniques etc.	After this section, participants are able to understand Analytics Spectrums, Data Science Lifecycle, Commonly used data sciences techniques	2.0
2	Day1 [11.15am-12.45pm] - Introduction Machine Learning and MLLib	This section covers Machine Learning and Top 5 MLLib	After this section, participants are able to understand Machine Learning and Top 5 MLLib	1.5

3	Day1 [1.45pm-3.45pm] - Data Acquisition and Cleansing	This section covers Data Acquisition and Cleansing	After this section, participants are able to perform Data Acquisition and Cleansing	2.0
4	Day1 [4pm-5.30pm] - Exploratory Data Analysis	This section covers Exploratory Data Analysis and methods	After this section, participants are able to understand Exploratory Data Analysis and methods	1.5
5	Day2 [9am-11am] - Types of Machine Learning and Regression models	This section covers Types of Machine Learning and Regression models	After this section, participants are able to understand Types of Machine learning and Regression models	2.0
6	Day2 [11.15am-12.45pm] - Regression models - Performance Evaluation for Regression models	This section covers Regression models and Performance Evaluation	After this section, participants are able to understand Regression models and perform Performance Evaluation	1.5
7	Day2 [1.45pm-3.45pm] - Classification models - Performance Evaluation for Classification models	This section covers Classification models and Performance Evaluation	After this section, participants are able to understand Classification models and perform Performance Evaluation	2.0
8	Day2 [4pm-5.30pm] - Classification models - Performance Evaluation for Classification models (continue)	This section covers Classification models and Performance Evaluation	After this section, participants are able to understand Classification models and perform Performance Evaluation	1.5
9	Day3 [9am-11am] - Clustering models	This section covers Clustering models	After this section, participants are able to understand Clustering models	2.0
10	Day3 [11.15am-12.45pm] - Geo Clustering	This section covers Geo Clustering	After this section, participants are able to	1.5

			understand Geo Clustering	
11	Day3 [1.45pm-3.45pm] - Graph Analysis	This section covers Graph Analysis	After this section, participants are able to understand Graph Analysis	2.0
12	Day3 [4pm-5.30pm] - Recommender System	This section covers Recommender System Using Apriori Algorithm and other methods	After this section, participants are able to perform Recommender System Using Apriori Algorithm and other methods	1.5
13	Day4 [9am-11am] - Natural Language Processing (NLP)	This section covers Text Processing and NLP, the Important Libraries for NLP (using Python)	After this section, participants are able to understand Text Processing and the Important Libraries for NLP (using Python); perform NLP	2.0
14	Day4 [11.15am-12.45pm] - Sentiment Analysis	This section covers Sentiment Analysis	After this section, participants are able to perform Sentiment Analysis	1.5
15	Day4 [1.45pm-3.45pm] - Model Evaluation Metrics	This section covers Types of Model Evaluation Metrics, Test Sets, Learning Curve, Limitations of Single Training/Test Partition	After this section, participants are able to understand Types of Model Evaluation Metrics, Test Sets, Learning Curve, Limitations of Single Training/Test Partition	2.0
16	Day4 [4pm-5.30pm] - Model Evaluation Metrics (continue)	This section covers Stratified Sampling, Cross Validation, Confusion Matrix, Other Accuracy Metrics	After this section, participants are able to understand Stratified Sampling, Cross Validation, Confusion Matrix, Other Accuracy Metrics etc.	1.5