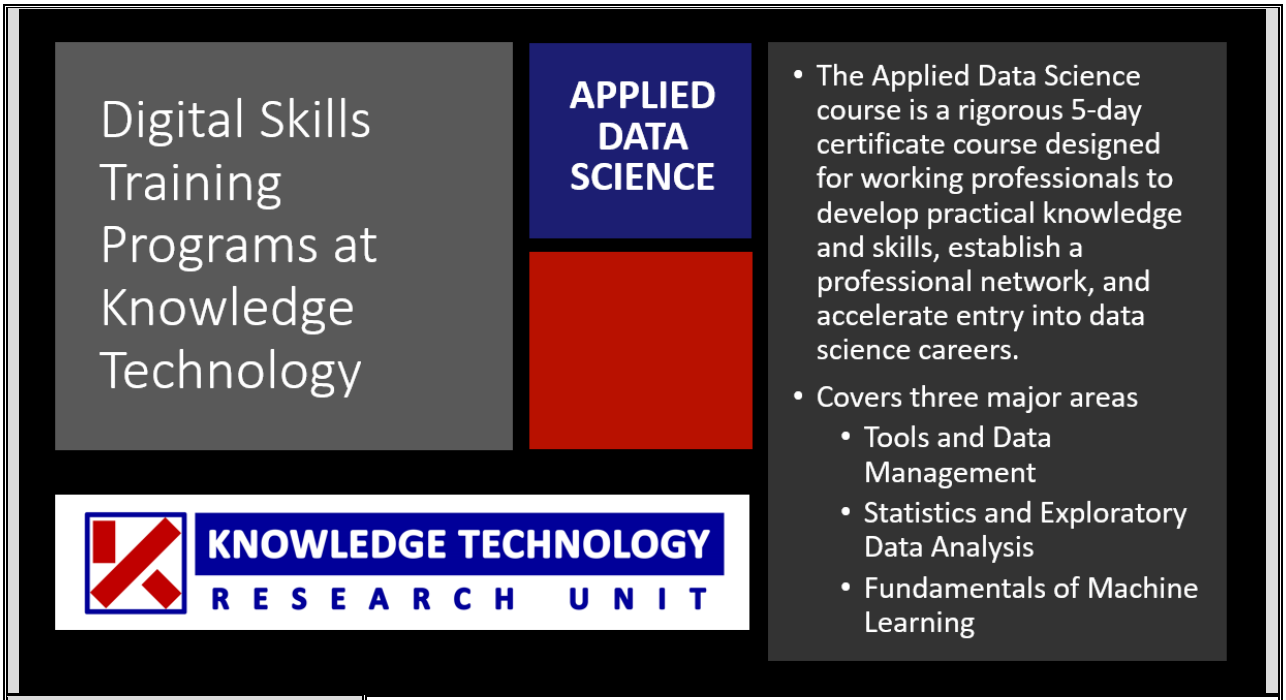


Course Title	Applied Data Science
Duration	5 Days
Trainer	Assoc. Prof. Dr. Rayner Alfred
Cost	Email ralfred121@gmail.com or call 013-881-9966 for quotations
Max Participants	25

SYNOPSIS

Our approach to this course is to teach the underlying the underlying concepts and statistics of Data Science. Going beyond the theory, our approach invites participants to go through several practical sessions, where learning is facilitated by live subject matter experts and enriched by practitioners in the field of Data Science.

LEARNING OUTCOMES

The course requires learners to work on application projects. These projects require learners to apply the Data Science concepts they have learned to datasets and derive inferences. These application projects are intentionally made to be challenging. We expect learners to spend substantial time and effort solving the application projects. At the end of the course, we expect learners to be able to apply Data Science methods to analyse data and to solve many of the business problems they face at their workplace.

JUSTIFICATION TO LEARN APPLIED DATA SCIENCE

Data has been called the new global currency, and its meteoric rise is transforming entire industries—and driving the demand for practitioners who can wield its power. From health care and finance to entertainment, cybersecurity and beyond, the need for data scientists continues to grow in tandem with opportunities for career advancement within the field. This Applied Data Science course is designed to help fill this talent gap and further the use of data science to solve real-world problems.

TOPICS LIST

Tools and Data Management

[1] Python Basics
[2] Intermediate Python
[3] Relational Databases
[4] Structured Query Language (SQL)

Statistics and Exploratory Data Analysis

[5] Statistical Distributions
[6] Sampling
[7] Hypothesis Testing
[8] Data Analysis and Visualization

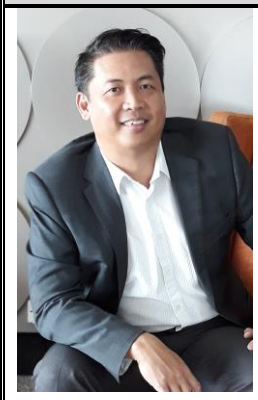
Fundamentals of Machine Learning

[9] Text Mining
[10] Regression and Classification
[11] Clustering and Decision Trees

COURSE SYLLABUS (5 DAYS)

DAY	TOPICS COVERED	TIME
One	**MODULE 1: PYTHON BASIC** ➢ Translating procedures into code	9:00 am – 12:30 pm
	MODULE 2: INTERMEDIATE PYTHON ➢ Introduction to Data Structures	2:30 pm – 4:00 pm
Two	**MODULE 3: RELATIONAL DATABASE** ➢ Where (most) data is stored	9:00 am – 12:30 pm
	MODULE 4: STRUCTURED QUERY LANGUAGE (SQL) ➢ Ubiquitous database formats/languages	2:30 pm – 4:00 pm
Three	**MODULE 5: STATISTICAL DISTRIBUTIONS** ➢ The shape of Data	9:00 am – 10:30 1m
	MODULE 6: SAMPLING ➢ When you don't have all the data	11:00 am – 12:30 pm
	MODULE 7: HYPOTHESIS TESTING ➢ Answering questions about your data	2:30 pm – 4:00 pm
Four	**MODULE 8: DATA ANALYSIS & VISUALIZATION** ➢ Using Python's NumPy for analysis	9:00 am – 12:30 pm
	MODULE 9: DATA ANALYSIS & VISUALIZATION ➢ Using Python's Pandas for data wrangling	2:30 pm – 4:00 pm
Five	**MODULE 10: TEXT MINING** ➢ Automatic understanding of text	9:00 am – 10:30 1m
	MODULE 11: REGRESSION AND CLASSIFICATION ➢ Machine learning methods for prediction	11:00 am – 12:30 pm
	MODULE 12: CLUSTERING AND DECISION TREES ➢ Machine learning methods for representation	2:30 pm – 4:00 pm

TRAINER'S BIOGRAPHIES



RAYNER ALFRED

ASSOCIATE PROFESSOR OF COMPUTER SCIENCE

Certified IBM DB2 Academic Associate, Certified Tester Foundation Level (CTFL)

AREAS OF SPECILIZATION: Advanced Machine Intelligence, Data Analytics, Data Mining, Information Retrieval, Artificial Intelligence, Machine Learning, Knowledge Discovery

ADDRESS: Knowledge Technology Research Group, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah.

CONTACT: Mobile: 6013-881-9966, eMail: ralfred@ums.edu.my

Rayner Alfred is an Associate Professor of Computer Science at the Faculty of Computing and Informatics, Universiti Malaysia Sabah in Malaysia that focuses on Data Science and Software Engineering programmes. He leads and defines projects around knowledge discovery, information retrieval and machine learning that focuses on building smarter mechanism that enables knowledge discovery in structured and unstructured data. His work addresses the challenges related to big data problem: How can we create and apply smarter collaborative knowledge discovery and machine learning technologies that bridge the structured and unstructured data mining and cope with the big data problem.

Rayner completed his PhD in 2008 looking at intelligent techniques using machine learning to model and optimize the dynamic and distributed processes of knowledge discovery for structured and unstructured data. He holds a PhD degree in Computer Science from York University (United Kingdom), a master's degree in computer science from Western Michigan University, Kalamazoo (USA) and a Computer Science degree from Polytechnic University of Brooklyn, New York (USA) where he was the recipient of the *Myron M. Rosenthal Academic Achievement Award* for the outstanding academic achievement in Computer Science in 1994. He has authored and co-authored more than 100 journals/book chapters and conference papers, editorials, and served on the program and organizing committees of numerous national and international conferences and workshops.

Rayner is currently a member of IEEE, a Certified Software Tester (CTFL) from the International Software Testing Qualifications Board (*ISTQB*), and a certified IBM DB2 Academic Associate (IBM DB2 AA). He leads the Advanced Machine Intelligence (AMI) research group in UMS and he has led several projects related to knowledge discovery and machine learning on Big Data. Rayner is also the recipient of the Research Fellow at Japan Advanced Institute of Science and Technology (JAIST), Japan. He is also the recipient of multiple GOLD awards at national and international research exhibitions in Data Mining and Machine Learning based solutions (Face Recognition and Knowledge Discovery), that include International Trade Fair Ideas in Nuremberg, Germany (iNEA2018) International Invention Innovation Competition in Toronto, Canada (iCAN 2018), Seoul International Invention Exhibition in Seoul, Korea (SIIF 2010). He has secured RM6,931.433.00 worth of project grants. Some of his project researches include biometric authentication using face recognition, building security based on plate number recognition using deep learning, sentiment analysis for Malay and English in measuring public opinion, news-news correlation trending, machine learning algorithm-based solution for predicting diseases in health care, smart monitoring using an ensemble based face recognition system and smart information management and retrieval to name a few.