

# PERANTIS MySTEP Data Analyst

1st July – 31st December 2022

A Malaysia short-term employment programme supporting the Malaysian Government's initiative to revive the economy via human capital development





## Foreword

Perantis MySTEP is a Malaysia short-term employment programme supporting the Malaysian government initiative to revive the economy via human capital development

KOTO Association is an Artificial Intelligence-driven platform for Upskilling, Job Matching and Financial Inclusion. KOTO Association is honoured to be designated as the Authorized Training Partner (ATP) for this programme.

The outcome of the training programme is to enhance Iskandar Malaysia's regional economic development via human capital development, produce competent talent for employment and address the unemployment issues caused by the pandemic Covid-19.

To actualize this initiative, Perantis Iskandar has commissioned its Authorized Industry Partner (AIP), My Xpress Train Sdn. Bhd., to train trainees on practicing their knowledge during industrial training.

KOTO Association will deliver a 1-month classroom training in different categories as stated below, followed by 5 months of industrial training after which the trainees may begin full-time employment at My Xpress Train Sdn. Bhd based on their performance during the programme.

It is a pleasure to welcome all the trainees to this programme. Do take advantage of the good opportunity to kick-start your career with a better foundation.

Appreciation goes to Perantis Iskandar and the Malaysian Government for this initiative and the opportunity to be the Authorized Training Partner (ATP).

Thank you.  
Steve Edwards

C  
O  
U  
R  
S  
E  
  
D  
A  
T  
E  
  
2  
0  
2  
2

**Data Engineering**  
4<sup>th</sup> July – 8<sup>th</sup> July

- Intro to Data Engineering
- Intro to Hadoop Ecosystem (Hadoop DFS, Spark, Kafka, HBase)
- Case Study for Hadoop
- Database in SQL/noSQL in DE
- ETL Pipeline

**Artificial Intelligence**  
18<sup>th</sup> July – 22<sup>nd</sup> July

- AI Nature & Problem Definition
- Uninformed/ Informed Search
- Ontology, Classification, Semantic Network
- Python – Lecture Slide
- Natural Language Processing

**Data Science**  
11<sup>th</sup> July – 15<sup>th</sup> July

- Introduction to Data Science
- Data Science Process
- Visualization and Descriptive Analysis
- Mining Data Streams
- Case Study

**Machine Learning**  
25<sup>th</sup> July – 29<sup>th</sup> July

- Supervised Learning
- Unsupervised Learning
- Neural Network
- Image Processing and Computer Vision

# Data Engineering

Date: 4<sup>th</sup> July – 8<sup>th</sup> July 2022

Trainer:  
Prof Dr Lim Tong Ming

Mr. Tan Zi Xuan

## Course Description

Data Engineering or Information Engineering is a software approach to develop information. Good data engineering skills help data engineers to optimize data of a company towards usability. Trainees will learn about the Fundamentals of Data Engineering, Fundamentals of Hadoop Systems, Database in SQL/noSQL in Data Engineering and ETL Pipeline. With this course, trainees will have better understanding of data engineering and able to solve the problem of data storage.

## Course Schedule

| Day & Date                             | Mode                               | Time                    | Activities                                      | Trainers               |
|--|------------------------------------|-------------------------|---|------------------------|
| Monday<br>4 <sup>th</sup> July 2022    | Physical Lesson (The Xpress Train) | 9:00 AM                 | Lecture, Exercise                               | Mr. Tan Zi Xuan        |
| Tuesday<br>5 <sup>th</sup> July 2022   |                                    | –<br>4:00 PM            |   |                        |
| Wednesday<br>6 <sup>th</sup> July 2022 | Hybrid (Google Meet)               | 10:00 AM                | Online Lecture (2 hours),<br>Exercise (4 hours) | Prof Dr. Lim Tong Ming |
| Thursday<br>7 <sup>th</sup> July 2022  |                                    | –<br>4:00 PM            |   |                        |
| Friday<br>8 <sup>th</sup> July 2022    | Google Chat                        | Subject to each trainer | Assignment                                      |                        |
| Saturday<br>9 <sup>th</sup> July 2022  | <b>Rest</b>                        |                         |   |                        |
| Sunday<br>10 <sup>th</sup> July 2022   |                                    |                         |   |                        |

# Data Science

Date: 11<sup>th</sup> July – 15<sup>th</sup> July 2022

Trainer:  
Dr Sean Tan Chi Wee

## Course Description

Data Science frequently used to make the decisions or predictions of predictive causal analytics, prescriptive analytics, and machine learning. Data Science is the future of Artificial Intelligence as it can help on processing the data. During this Data Science course, trainees will learn about the Fundamentals of Data Science, Data Science Process, Visualization and Descriptive Analysis, and Mining Data Streams. The case study will be given to trainees to allowed them to have more deeper understanding of data science.

## Course Schedule

| Day & Date                              | Mode                               | Time                       | Activities                                      | Trainer              |
|---|------------------------------------|----------------------------|---|----------------------|
| Monday<br>11 <sup>th</sup> July 2022    | Physical Lesson (The Xpress Train) | 9:00 AM<br>–<br>4:00 PM    | Lecture, Exercise                               | Dr. Sean Tan Chi Wee |
| Tuesday<br>12 <sup>th</sup> July 2022   |                                    |                            |   |                      |
| Wednesday<br>13 <sup>th</sup> July 2022 | Hybrid (Google Meet)               | 10:00 AM<br>–<br>4:00 PM   | Online Lecture (2 hours),<br>Exercise (4 hours) |                      |
| Thursday<br>14 <sup>th</sup> July 2022  |                                    |                            |   |                      |
| Friday<br>15 <sup>th</sup> July 2022    | Google Chat                        | Subject to<br>each trainer | Assignment                                      |                      |
| Saturday<br>16 <sup>th</sup> July 2022  | <b>Rest</b>                        |                            |   |                      |
| Sunday<br>17 <sup>th</sup> July 2022    |                                    |                            |   |                      |

# Artificial Intelligence

Date: 18<sup>th</sup> July – 22<sup>nd</sup> July 2022

Trainer:

Dr. Sean Tan Chi Wee

## Course Description

This Artificial Intelligence course allowed the trainees to know how they can use advanced Mathematics, Computer Science, and Engineering to create a system or a machine. In today's world, Artificial Intelligence acts as the main driver of emerging technologies. So, the knowledge of Artificial Intelligence can increase the value of trainees who would like to work in this industry in the future. AI Nature and Problem Definition, Uninformed or Informed Search, Ontology, Classification, Semantic Network, Python, Natural Language Processing are the contents that will be teach by trainer in this course.

## Course Schedule

| Day & Date                              | Mode                               | Time                       | Activities                                      | Trainer              |
|---|------------------------------------|----------------------------|---|----------------------|
| Monday<br>18 <sup>th</sup> July 2022    | Physical Lesson (The Xpress Train) | 9:00 AM<br>–<br>4:00 PM    | Lecture, Exercise                               | Dr. Sean Tan Chi Wee |
| Tuesday<br>19 <sup>th</sup> July 2022   |                                    |                            |   |                      |
| Wednesday<br>20 <sup>th</sup> July 2022 | Hybrid (Google Meet)               | 10:00 AM<br>–<br>4:00 PM   | Online Lecture (2 hours),<br>Exercise (4 hours) |                      |
| Thursday<br>21 <sup>st</sup> July 2022  |                                    |                            |   |                      |
| Friday<br>22 <sup>nd</sup> July 2022    | Google Chat                        | Subject to<br>each trainer | Assignment                                      |                      |
| Saturday<br>23 <sup>rd</sup> July 2022  | <b>Rest</b>                        |                            |   |                      |
| Sunday<br>24 <sup>th</sup> July 2022    |                                    |                            |   |                      |

# Machine Learning

Date: 25<sup>th</sup> July – 26<sup>th</sup> July 2022

Trainer:

Assoc Prof Ts. Dr. Tew Yiqi

## Course Description

Machine Learning is a must-have skill for data analyst, and AI programmer. It helps data analyst to transform all the unprocessed data into accurate trends and forecasts. Through this course, trainees will learn and practice the machine learning skills. Supervised Learning, Unsupervised Learning, Neural Network, Image Processing and Computer Vision are the content that will be teach by trainer in this course.

## Course Schedule

| Day & Date                              | Mode                               | Time                       | Activities                                      | Trainer                     |
|---|------------------------------------|----------------------------|---|-----------------------------|
| Monday<br>25 <sup>th</sup> July 2022    | Physical Lesson (The Xpress Train) | 9:00 AM<br>–<br>4:00 PM    | Lecture, Exercise                               | Assoc Prof Ts. Dr. Tew Yiqi |
| Tuesday<br>26 <sup>th</sup> July 2022   |                                    |                            |   |                             |
| Wednesday<br>27 <sup>th</sup> July 2022 | Hybrid (Google Meet)               | 10:00 AM<br>-<br>4:00 PM   | Online Lecture (2 hours),<br>Exercise (4 hours) |                             |
| Thursday<br>28 <sup>th</sup> July 2022  |                                    |                            |   |                             |
| Friday<br>29 <sup>th</sup> July 2022    | Google Chat                        | Subject to<br>each trainer | Assignment                                      |                             |
| Saturday<br>30 <sup>th</sup> July 2022  | <b>Rest</b>                        |                            |   |                             |
| Sunday<br>31 <sup>st</sup> July 2022    |                                    |                            |   |                             |

## TRAINERS



### Prof Dr Lim Tong Ming

Prof. Dr Lim Tong Ming is the director of Centre for Business Incubation and Entrepreneurial Ventures (CBIEV), Tunku Abdul Rahman University College (TARC), Malaysia. He has 10 years experiences of commercial software and almost 20 years of academic experience. He is a researcher that focused the areas from object technologies on software engineering to databases, peer to peer technologies, knowledge sharing to social media analytics and social influence maximization. He is also a professional ACM member and a society affiliate member of IEEE Association.



### Assoc Prof Ts Dr Tew Yiqi

Assoc Prof Ts Dr Tew Yiqi is a senior lecturer from the Faculty of Computing & Information Technology (FOCS), Tunku Abdul Rahman University College (TARC), Malaysia. He has more than 10 years experiences of working in the higher education and smart home industry. He is a lead researcher in Computational Intelligence with publication more than 30 papers including ISI Tier 1 papers. He also engaged the industrial projects on Industry 4.0, Data Channeling work on Smart Manufacturing, Smart Agricultural health status monitoring and Smart Campus initiatives.



### Dr Sean Tan Chi Wee

Dr Sean Tan Chi Wee is an enthusiastic researcher in conducting and supporting research into Image Processing. He has more than 10 years of education and hands-on experience in teaching. He is also adaptive team player with expertise in data collection, confidential document control and problem-solving to complete projects on time.



### Mr. Tan Zi Xuan

Mr. Tan Zi Xuan is a master's student at Tunku Abdul Rahman University College (TARC), Malaysia. He has a strong background in computer science such as programming, database, data science, and tools/ software. He has work experience on administer and managing the TARUC Hadoop Cluster.