

RTTP APPLICATION  
AVAILABLE

DATA SCIENTIST DEVELOPMENT PROGRAMME

# MACHINE LEARNING AI MODULE

## MODULE

01

MACHINE LEARNING -  
FUNDAMENTALS

02

MACHINE LEARNING -  
REGRESSION ANALYSIS

03

MACHINE LEARNING MODELS -  
DECISION TREE AND  
RANDOM FOREST CLASSIFIERS

04

MACHINE LEARNING MODELS -  
CLUSTER ANALYSIS AND  
ARTIFICIAL NEURAL NETWORK

## COURSE DETAILS

COURSE FEE : ~~\$16,000~~ **\$5,333**

HOURS : 24

LANGUAGE : CANTONESE

THIS COURSE IS APPROVED BY REINDUSTRIALISATION AND TECHNOLOGY TRAINING PROGRAMME (RTTP), WHICH OFFERS UP TO **2/3 COURSE FEE** REIMBURSEMENT UPON SUCCESSFUL APPLICATION.

COMPANIES SHOULD SUBMIT THEIR RTTP TRAINING GRANT APPLICATION FOR THEIR EMPLOYEE(S) VIA [HTTPS://RTTP.VTC.EDU.HK/RTTP/LOGIN](https://rttp.vtc.edu.hk/rttp/login) **AT LEAST TWO WEEKS** BEFORE COURSE COMMENCEMENT.



COURSE REGISTRATION



RTTP APPLICATION

## SELECTIVE SPECIALTY

THERE ARE VARIOUS TYPES OF FOCUS ON LEARNING DATA SCIENCE:



### MANAGEMENT USERS

- Focus on data visualization, general understanding of AI and Big Data analysis



### JUNIOR COURSE TAKERS

- Are more interested in learning programming skills



### DATA ANALYSTS

- Favor machine model building and statistical skills



### ALL-ROUNDERS

- Have a larger appetite and wider spectrum of interest



### BUSINESS INTELLIGENCE DEVELOPERS

- Have a larger appetite and wider spectrum of interest

HAVING UNDERSTOOD THE OWN ABILITY,  
YOU ARE **OPTIONAL** TO TAKE CERTAIN MODULES TO MEET YOUR NEEDS AND SATISFY YOUR INTEREST.

## PRE-REQUISITES

WE ASSUME THAT THE COURSE TAKERS HAVE THE KNOWLEDGE ON OR EQUIVALENT TO :



**01** EXCEL (TABLES AND PIVOT TABLES)



**02** BASIC STATISTICS

## TRAINING METHODS

THE TEACHING OF THIS COURSE CONSISTS OF THE FOLLOWING ELEMENTS :



LECTURES & DEMONSTRATION



INTERACTIVE CLASS DISCUSSION



CASE STUDIES

**START YOUR DATA SCIENCE JOURNEY NOW !**

# ML1 MACHINE LEARNING - FUNDAMENTALS

## INTRODUCTION TO AI, MACHINE LEARNING AND DEEP LEARNING :

- 1 MACHINE LEARNING ALGORITHMS
- 2 MACHINE LEARNING MODEL BUILDING PROCESS AND PROJECT CHECKLIST
- 3 TYPES OF MACHINE LEARNING MODELS
- 4 ASSESSMENT OF MACHINE LEARNING MODELS
- 5 PREDICTIVE ANALYTICS – STATISTICAL APPROACH VS MACHINE LEARNING APPROACH
- 6 INTRODUCTION TO DATA SCIENCE PROGRAMMING
- 7 DATA CLEANSING TECHNIQUES

## CASE STUDY

- 01** HOW CAN WE DOWNLOAD FINANCIAL DATA[\*] FROM YAHOO AND VISUALIZE THE PRICE TREND WITH AASTOCK-STYLE DIAGRAMS FOR STATISTICAL ANALYSIS (PROGRAMMING APPROACH)?

### [\*] FINANCIAL DATA AND INDICATORS COVER :

- HK /US STOCK PRICES, FOREIGN EXCHANGE RATE, GOLD PRICE AND PLATINUM PRICE
- STOCK TURNOVER, MOVING AVERAGE (SMA/WMA), BOLLINGER BAND, RSI, MACD, RATE OF RETURN, PRICE HISTOGRAM, ETC.



# ML2 MACHINE LEARNING - REGRESSION ANALYSIS

## ILLUSTRATION OF SUPERVISED MACHINE LEARNING ALGORITHMS :

- 1 MULTIPLE LINEAR REGRESSION
- 2 LOGISTIC REGRESSION
- 3 MULTINOMIAL REGRESSION
- 4 ANALYSIS ON THE RELATIVE IMPORTANCE OF VARIABLES

## CASE STUDY

- 01** HOW CAN WE PREDICT MACHINERY EXPENSE WITH HISTORICAL RECORDS? (MULTIPLE LINEAR REGRESSION)
- 02** HOW CAN WE MAKE STOCK BUYING AND SELLING DECISIONS WITH PAST RECORDS? (LOGISTIC REGRESSION)



# ML3

## MACHINE LEARNING - DECISION TREE AND RANDOM FOREST CLASSIFIERS

### MORE ON SUPERVISED MACHINE LEARNING ALGORITHMS :

- 1 DECISION TREE (CART - CLASSIFICATION & REGRESSION TREE)
- 2 ENTROPY, INFORMATION GAIN AND GINI INDEX
- 3 DECISION TREE PRUNING
- 4 MODEL ENSEMBLE (BAGGING, BOOSTING, SUBSPACE SAMPLING)
- 5 RANDOM FOREST

### CASE STUDY

- 01 HOW DO WE DETERMINE/PREDICT WHETHER A CASE OF BREAST CANCER IS BENIGN OR MALICIOUS FROM THE PAST RECORDS? (DECISION TREE)
- 02 HOW CAN WE PERFORM SENTIMENTAL ANALYSIS BASED ON TEXTUAL CUSTOMER FEEDBACK COLLECTED FROM A WEBSITE FOR A RESTAURANT OWNER? (RANDOM FOREST)

# ML4

## MACHINE LEARNING - CLUSTER ANALYSIS AND ARTIFICIAL NEURAL NETWORK

### CLUSTER ANALYSIS

- 1 K-MEANS CLUSTERING
- 2 HIERARCHICAL AGGLOMERATIVE CLUSTERING
- 3 PARTITIONING AROUND MEDOIDS

### DEEP LEARNING ALGORITHM

- 1 CONCEPTS OF DL AND CASES OF PROBLEMS
- 2 ARTIFICIAL NEURAL NETWORK (ANN)
- 3 STOCHASTIC GRADIENT DESCENT (SGD), MINI-BATCH AND STANDARD GRADIENT DESCENT
- 4 VISUALIZATION OF MODEL ASSESSMENT ON ANN PROCESSING

### CASE STUDY

- 01 WHAT IS THE MOST POWERFUL CLUSTERING ALGORITHM?  
DBSCAN : DENSITY-BASED SPATIAL OF APPLICATION WITH NOISE
- 02 HOW DO WE CATEGORIZE WINES BASED ON THE INGREDIENTS AND COMPARE THE PERFORMANCE WITH THE EXPERT CLASSIFICATION?
- 03 HOW CAN WE MAKE STOCK BUYING AND SELLING DECISION WITH ARTIFICIAL NEURAL NETWORK? (VS THE PREVIOUS LOGISTICS REGRESSION APPROACH) NOISE
- 04 WHY IS A COMPUTER ABLE TO UNDERSTAND HAND-WRITTEN DIGITS ?

