Professional Certificate in Data Science and Analytics - Final Assessment.

Total points 43/50

Please read the question and select the right answer.

0 of 0 points

Name *

YONG ZONG WEI

Contact Number: *

0189107110

Email: *

soyayong@gmail.com

IC Number: *

930410055209

Please read the questions and select the right answer. 43 of 50 points

✓ Data Science is the field of _____ * 1/1

- Exploring data
- Manipulating Data
- Analysing data
- (✓) All of the above

✓ Which of the below statement is correct? * 1/1

- To be successful data scientist working on big data you must have a degree in engineering or statistic
- (✓) Data Scientist/Big Data is not a discipline study taught at universities, everyone can be a Data Scientist as long as they have essential characteristics.

✓ The important characteristics of being a Data Scientist working on Big Data are: * 1/1

- Curious
- Judgmental
- Argumentative and able to tell a story
- (✓) All of the above

✓ There are many paths to a career in data science; most, but not all, involve a little math, a little science, and a lot of curiosity about data. * 1/1

- (✓) True
- False

✓ Which is the of the below statement is correct about Cloud for data science? * 1/1

- (✓) It allows you to pass the limitations of your personal machine (Laptop/Desktop)
- Has the limitation of not being able to deploy.

✓ What are the V's of Big Data? * 1/1

- ○ Velocity
- ○ Volume & Variety
- ○ Veracity & Value
- ◉ All of the above ✓

✓ Hadoop is a collection of open-source software utilities using a network of many computers to solve problems involving Big Data. * 1/1

- ◉ True ✓
- ○ False

✓ Data Science is a part of Data Mining. * 1/1

- ○ True
- ◉ False ✓

✓ Data Science Methodology has 10 stages, from Business Understanding to Feedback. * 1/1

- ◉ True ✓
- ○ False

✓ Business Understanding in the methodology of Data Science which is about understanding the problem and defining the main goal and objective. * 1/1

- ◉ True ✓
- ○ False

✓ Analytics Approach is the second stage in the methodology of Data Science, it bases on the first stage, and it's about finding the best algorithm to solve the problem. * 1/1

- ◉ True ✓
- ○ False

✗ Data Collection stage is about checking the availability of required data. * 0/1

- ○ True
- ◉ False ✗

Correct answer

- ◉ True

✗ Data Understanding takes almost 70% - 90% of the time of the Data Science Project * 0/1

- ◉ True ✗
- ○ False

Correct answer

- ◉ False

✓ Data Preparation is divided into: * 1/1

- ○ Data Cleaning/Data Wrangling
- ○ EDA
- ◉ Both of the above ✓

✓ A predictive model can be built during the Modelling Stage * 1/1

- ◉ True ✓
- ○ False

✓ F1-score is based on Confusion Matrix, and it is consider one of the method of the Evaluation Stage. * 1/1

- (●) True ✓
- ( ) False

✓ The methodology of Data Science is iterative * 1/1

- (●) True ✓
- ( ) False

✓ Using the skill network lab or Colab does not need to install any software on your machine, and you can store you coding on cloud(your account). * 1/1

- (●) True ✓
- ( ) False

✓ Python is object oriented multi-purposes computer programming language widely used among data scientist. * 1/1

- (●) True ✓
- ( ) False

✓ Python has integer, float, and string data types, concatenate any of those data types are accepted (like string + float). * 1/1

- ( ) True
- (●) False ✓

✓ Tuple is immutable while list is mutable. * 1/1

- (●) True ✓
- ( ) False

✗ The variable release_year_dict is a Python Dictionary, the result of applying the following method: release_year_dict.keys(), is to retrieve the values of the keys. (True, False). * 0/1

- ( ) True
- (●) False ✗

Correct answer

- (●) True

✓ Data Sets format can be in: * 1/1

- ( ) CSV
- ( ) Json
- ( ) Excel and SQL
- (●) All of the above ✓

✓ Python library is a collection of Libraries, functions, and methods that allow you to perform lots of actions without writing a code from scratch. * 1/1

- (●) True ✓
- ( ) False

✓ Scientific Python libraries are like Pandas, Numpy, and Scipy * 1/1

- (●) True ✓
- ( ) False

✓ Visualizations libraries are like Matplotlib, and Seaborn. * 1/1

- [x] True
- [ ] False

✓ Algorithmic libraries are like Scikit-learn, and Statsmodels * 1/1

- [x] True
- [ ] False

✓ Pandas datatypes are: * 1/1

- [ ] Object
- [ ] Float64
- [ ] Datetime64
- [ ] Int64
- [x] All of the above

✓ Data Wrangling is about: * 1/1

- [ ] Dealing with the missing values
- [ ] Data formatting and normalization
- [ ] Data binning and turning categorical values to numeric
- [x] All of the above

✓ Dealing with the missing values can be done by replacing it with the mean/average value or the most frequent value. * 1/1

- [x] True
- [ ] False

✓ Data normalization approaches are: * 1/1

- [ ] Simple Feature Scaling
- [ ] Min-Max
- [ ] Z-score
- [x] All of the above

✓ Correlation and correlation-statistics are considered EDA in the stage of Data Preparation * 1/1

- [x] True
- [ ] False

✓ Positive correlation and negative correlation have the same significance importance. * 1/1

- [x] True
- [ ] False

✗ Which of the following statements are correct? * 0/1

- [ ] Machine learning is a field of Artificial Intelligence, and Deep Learning is a subfield of Machine Learning.
- [x] Artificial Intelligence is a field of Machine learning, and Deep Learning is a subfield of Machine Learning. ✗
- [ ] Deep Learning is a field of Machine learning, and Artificial Intelligence is a subfield of Machine Learning.
- [ ] None of the above

Correct answer

- [x] Machine learning is a field of Artificial Intelligence, and Deep Learning is a subfield of Machine Learning.

✓ Machine Learning types are: * 1/1

- [ ] Supervised Learning & Un-supervised Learning

- Semi-supervised Learning
- Reinforcement Learning
- ✓ All of above

✓ If a dataset comes with labels and the problem is about predicting the new data label, then to solve this issue we can use: * 1/1

- ✓ Supervised Learning
- Semi-supervised Learning
- Reinforcement Learning
- Un-supervised Learning

✓ If a dataset needs to be divided into categories or clusters, then we should use: * 1/1

- Supervised Learning
- Semi-supervised Learning
- Reinforcement Learning
- ✓ Un-supervised Learning

✓ Regression and classification algorithms are for supervised machine learning while clustering algorithms are for un-supervised machine learning? * 1/1

- ✓ True
- False

✗ Regression algorithms are about predicating a value based on continuously increased or decreased values. * 0/1

- True
- ✗ False

Correct answer

- True

✓ Single linear/non-linear regression using one independent variable while Multiple linear/no-linear regression using more than one independent variables. * 1/1

- ✓ True
- False

✓ Train and Test / Split evaluation approach is important to have high out of sample accuracy * 1/1

- ✓ True
- False

✗ Reinforcement learning is about building a model learns from the mistakes * 0/1

- True
- ✗ False

Correct answer

- True

✓ Semi supervised learning model has very few data labelled * 1/1

- ✓ True
- False

✓ Supervised machine learning has more evaluation methods in compare to un-supervised machine learning. * 1/1

- ✓ True
- False

✓ Recommender system tries to capture people' behaviour to predict what else they might like or want. * 1/1

- (●) True ✓
- ( ) False

✓ Content-based recommendation system recommends items to the users based on their profile. * 1/1

- (●) True ✓
- ( ) False

✗ Machine learning perform much better in comparing to Deep learning when the data increase * 0/1

- (●) True ✗
- ( ) False

Correct answer

- (●) False

✓ Choose the right sentence * 1/1

- (●) Deep learning algorithm doesn't need a feature engineering, it is capable of automatic feature engineering through its neural network. ✓
- ( ) Machine learning algorithm doesn't need a feature engineering, it is capable of automatic feature engineering through its algorithms.

✓ Self-driving car, automatic machine translation, automatically adding sound to silent movies are some of the deep learning application. * 1/1

- (●) True ✓
- ( ) False

✓ Some of the basic SQL commands are * 1/1

- ( ) Create Table
- ( ) Insert
- ( ) Update
- ( ) Delete
- ( ) Select
- (●) All of the above ✓