

BINUS University

Academic Career: <i>Undergraduate / Master / Doctoral *)</i>		Class Program: <i>International/Regular/Smart Program/Global Class*)</i>	
<input checked="" type="checkbox"/> Mid Exam <input type="checkbox"/> Final Exam <input type="checkbox"/> Short Term Exam <input type="checkbox"/> Others Exam : _____		Term : Odd /Even/ Short *)	
<input checked="" type="checkbox"/> Kemanggisan <input checked="" type="checkbox"/> Alam Sutera <input type="checkbox"/> Bekasi <input type="checkbox"/> Senayan <input type="checkbox"/> Bandung <input type="checkbox"/> Malang		Academic Year : 2020 / 2021	
Faculty / Dept. : School of Computer Science		Deadline	Day / Date : Jumat / 30 Apr 2021 Time : 13:00
Code - Course : COMP6577 - Machine Learning		Class : All Classes	
Lecturer : Team		Exam Type : Online	
*) <i>Strikethrough the unnecessary items</i>			
<i>The penalty for CHEATING is DROP OUT!!!</i>			

Learning Outcomes:

LO 1 : explain the fundamental of machine learning concept

LO 2 : interpret the distribution of dataset using regression method

LO 3 : experiment classification and clustering algorithm from given dataset

Seluruh pertanyaan pada ujian ini memiliki beberapa cara atau pendekatan dalam menjawabnya. Silahkan gunakan cara atau pendekatan yang anda ketahui untuk menyelesaikan pertanyaan dibawah ini. Anda dapat menggunakan *library python* apapun yang tersedia untuk membantu jawaban anda. Bagi mahasiswa/i yang terbukti mencontek, maka akan dikenakan sanksi yang berlaku.

I. Esai (10 poin)

Jelaskan pemahaman mengenai **Discrete Random Variables** dan **Continuous Random Variables** ! Berikan Contoh. (LO 1, 10 poin)

(Untuk 2-Angka terakhir NIM Mahasiswa Ganjil)

II. Kasus (90 poin)

1. [LO1, LO2, 20 poin]

Buka *file* data "Mall_Customers.csv" yang diberikan dan lakukan **Data Cleaning** dan **Visual Exploratory Data Analysis**, dengan detail yang dimaksud sebagai berikut:

- a. Siapkan Variabel yang akan dipakai *Spending Score, Genre, Age* dan *Annual Income*
- b. Periksa *missing values* dan *fill* NA dengan nilai rata-rata untuk *continuous variable*
- c. Visualisasikan dan jelaskan hubungan ≥ 2 variabel
- d. Periksa **Correlations** dari *independent variable* dengan *dependent variable*

2. [LO1, LO2, 20 poin]

Buatlah **Training Set** dan **Test Set**, lalu *print* **Training Set**.

3. [LO 1, LO 2, 20 poin]

Lakukan Pemodelan **Linear Regression**

Verified by,

Ajeng Wulandari (D6422) and sent to Program on Mar 30, 2021

- a. Prediksi Spending Score
 - b. Print Data Prediksi dan Aktual.
4. [LO 1, LO 2, 20 poin] Lakukan Evaluasi Pemodelan *Linear Regression* tersebut dengan cara berikut:
 - a. Evaluasi *Coefficient* untuk setiap *variable*
 - b. Evaluasi menggunakan *Metrics* (MAE,MSE, RMSE, dsb)
 5. [LO 1, LO 2, 10 poin] Lakukan **Plot Predicted Value vs Measure**

(Untuk 2-Angka terakhir NIM Mahasiswa Genap)

III. Kasus (90 poin)

1. [LO 1, LO 2, 20 poin]
Buka *file* data "titanic.txt" yang diberikan dan lakukan **Data Cleaning** dan **Visual Exploratory Data Analysis**, dengan detail yang dimaksud sebagai berikut:
 - a. Siapkan *Variable* yang akan dipakai *Survived*, *PassengerId*, *PClass*, *Age*, *SibSp* & *Parch*
 - b. Periksa *missing values* dan *fill* NA dengan nilai rata-rata untuk *continuous variable*
 - c. Visualisasikan dan jelaskan hubungan 2 atau 3 *variable*
 - d. Periksa **Correlations** dari *independent variable* dengan *dependent variable*
2. [LO 1, LO 2, 20 poin]
Buatlah **Training Set** dan **Test Set**, lalu print **Training Set**.
3. [LO 1, LO 2, 20 poin]
Lakukan Pemodelan **Linear Regression**
 - a. Prediksi Penumpang yang *Survived*
 - b. Print Data Prediksi dan Aktual.
4. [LO 1, LO 2, 20 poin]
Lakukan Evaluasi Pemodelan *Linear Regression* tersebut dengan cara berikut:
 - a. Evaluasi *Coefficient* untuk setiap *variable*
 - b. Evaluasi menggunakan *Metrics* (MAE,MSE, RMSE, dsb)
5. [LO 1, LO 2, 10 poin]
Lakukan **Plot Predicted Value vs Measure**

Kirimkan jawaban anda berupa a dan b:

- a. *File pdf* untuk jawaban yang berupa penjelasan dan hasil analisis **dan File source code python**, berupa *code* yang digunakan untuk mendapatkan penjelasan dan hasil analisis.
- b. Menggunakan **format Python Notebook** untuk menjelaskan secara detail dari hasil analisis yang disertai dengan *code* yang digunakan untuk mendapatkan penjelasan tersebut.

-- Selamat Mengerjakan --

Verified by,

Ajeng Wulandari (D6422) and sent to Program on Mar 30, 2021