

## SOAL TUGAS PENGGANTI PERTEMUAN

SEMESTER GANJIL 2020/2021

STMIK BANDUNG

<b>NAMA MATAKULIAH</b>	: Data Mining – Reg Pagi	<b>SKS</b>	: 3 SKS
<b>NAMA DOSEN</b>	: Siti Yuliyanti, ST., M.Kom	<b>HARI/ TANGGAL</b>	Deadline Sabtu, 14.00
<b>NIM</b>		<b>NAMA</b>	

1. Jelaskan mengapa data yang telah diambil/ di crawling perlu melalui tahapan data cleaning, data integration dan filtering dan jelaskan pula tujuan masing-masing tahapan tersebut! **(Nilai 20)**
2. Tentukan termasuk kedalam Class mana jika data terdiri dari Age : 33, Income : 270 menggunakan algoritma K-NN dengan K=4 ! **(Nilai 30)**

Age	Income	Class
29	350	A
51	430	B
33	290	A
24	255	A
40	410	B

3. Simulasikan menggunakan weka untuk dataset Surveillance.csv, gunakan klasifikasi naïve bayes dan menggunakan ALgoritma J48, bandingkan tingkat akurasi kedua algoritma tersebut terhadap datasets.cvs kemudian untuk tree tampilkan visualisasi treenya **(Nilai 50)!**

**Tampilkan screesshoot tree, dan setiap akurasi untuk masing-masing algoritma**