

KEAMANAN JARINGAN

Data Mining KNIME



OLEH :

Mochammad Jauhar Ulul Albab
(3122640044)

PROGRAM STUDI TEKNIK INFORMATIKA
DEPARTEMEN TEKNIK INFORMATIKA DAN
KOMPUTER POLITEKNIK ELEKTRONIKA
NEGERI SURABAYA

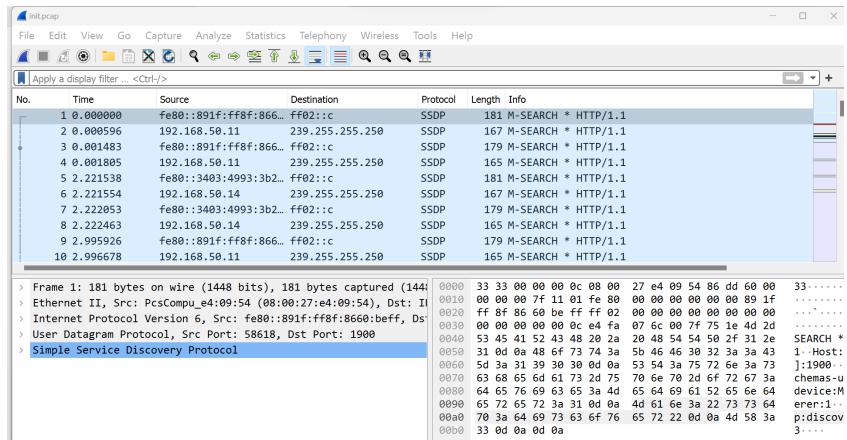
2022/2023

Laporan Praktikum

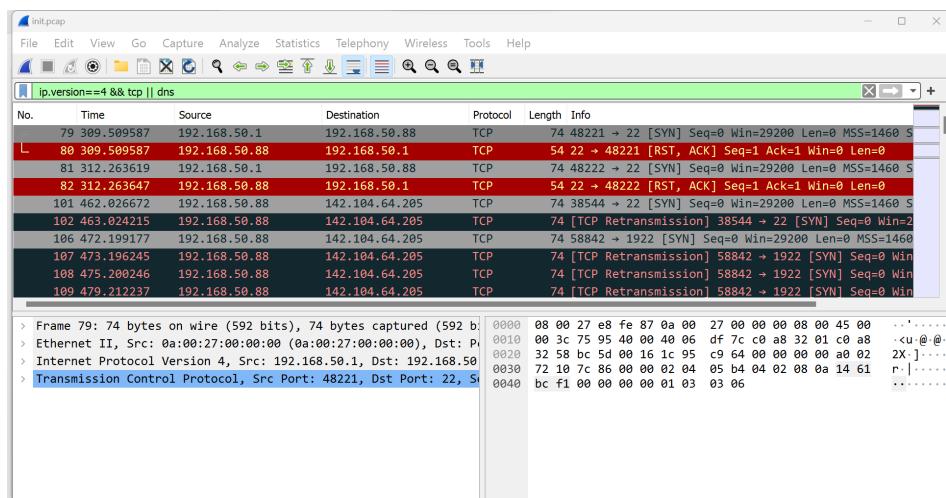
1. Terdapat 5 file hasil download yaitu, init.pcap, init2.pcap, init3.pcap, init4.pcap, init5.pcap

Name	Date modified	Type	Size
init	28/11/2017 16:52	Wireshark capture file	163.160 KB
init2	28/11/2017 16:51	Wireshark capture file	576.966 KB
init3	28/11/2017 16:51	Wireshark capture file	159.862 KB
init4	28/11/2017 16:52	Wireshark capture file	656.556 KB
init5	28/11/2017 16:52	Wireshark capture file	2 KB

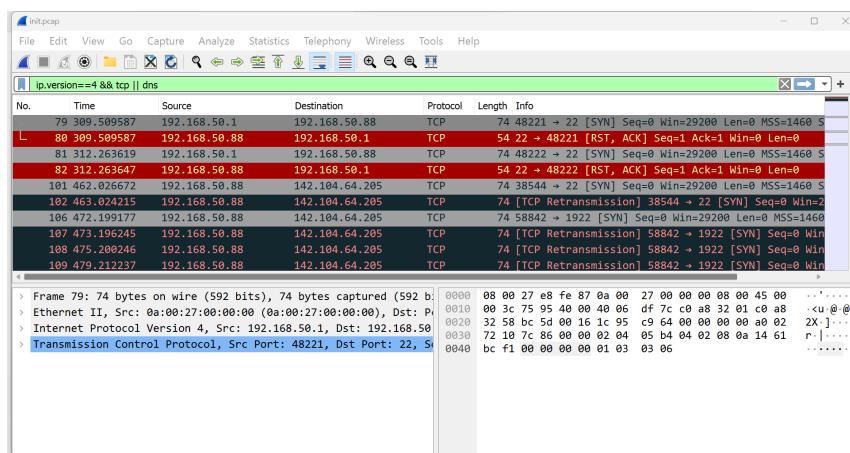
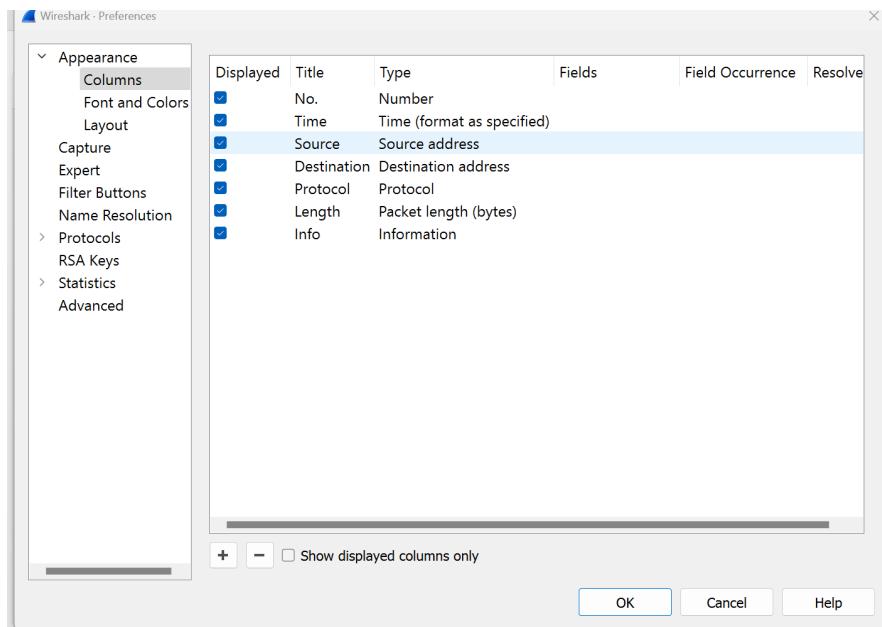
2. Kemudian buka file tersebut secara bergantian menggunakan Wireshark.
file init.pcap



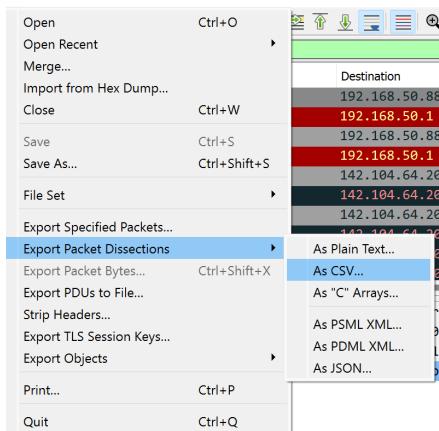
3. Untuk proses analisa yang akan dilakukan nantinya, kita mengambil data dengan ip versi 4 dan protocol TCP, DNS saja.



4. Untuk mendapatkan delta time display, klik Edit – Preferences – Column

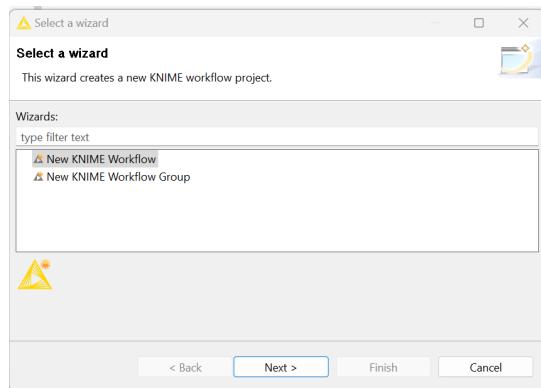


5. Export file pcap ke format csv.



Name	Date modified	Type	Size
init	08/06/2023 23:42	Microsoft Excel Com...	43.481 KB
init	28/11/2017 16:52	Wireshark capture file	163.160 KB
init2	08/06/2023 23:48	Microsoft Excel Com...	43.481 KB
init2	28/11/2017 16:51	Wireshark capture file	576.966 KB
init3	08/06/2023 23:51	Microsoft Excel Com...	149.525 KB
init3	28/11/2017 16:51	Wireshark capture file	159.862 KB
init4	09/06/2023 00:02	Microsoft Excel Com...	354.000 KB
init4	28/11/2017 16:52	Wireshark capture file	656.556 KB
init5	09/06/2023 00:25	Microsoft Excel Com...	1 KB
init5	28/11/2017 16:52	Wireshark capture file	2 KB

5. membuat workflow/project baru.



6. Tambahkan data kedalam file reader

Dialog - 3:1 - File Reader (Complex Format)

File

Settings Flow Variables Job Manager Selection Memory Policy

Enter ASCII data file location: (press 'Enter' to update preview)

Read from Local File System C:\Choirun Annas\JL D4 PENS\Keamanan Jaringan\Praktikum Data Mining\isot_app_and_botnet_dataset\botnet_data

Preserve user settings for new location Rescan

Basic Settings

read row IDs read column headers

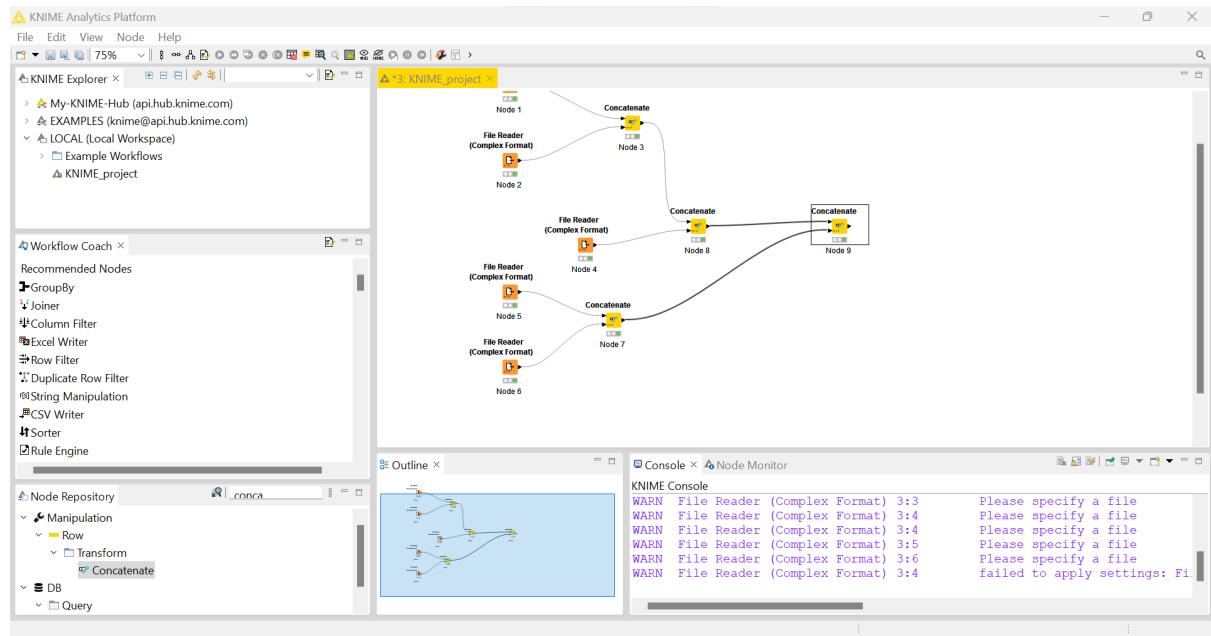
Column delimiter: , ignore spaces and tabs Java-style comments

Preview Click column header to change column properties (* = name/type user)

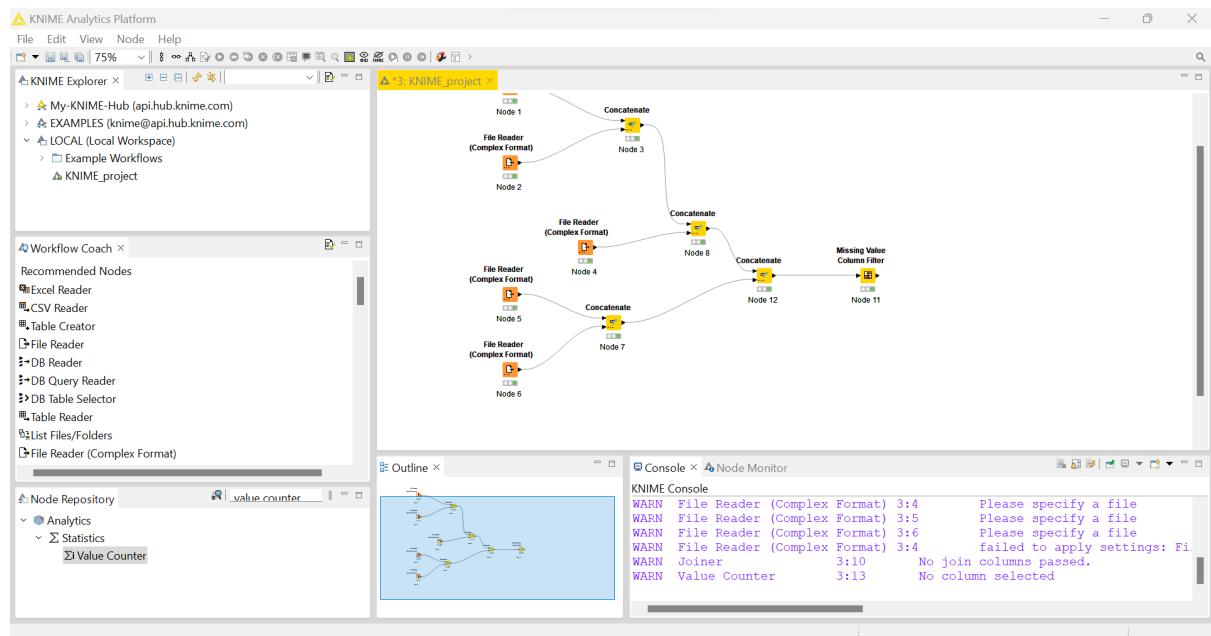
Row ID	D	Time	S	Source	S	Destinat...	S	Protocol	I	Length	S	Info
79		309.51		192.168.50.1		192.168.50.88		TCP	74		48221 > 22 [SYN] Seq=0 Win=29200	
80		309.51		192.168.50.88		192.168.50.1		TCP	54		22 > 48221 [RST, ACK] Seq=1 Ack=1	
81		312.264		192.168.50.1		192.168.50.88		TCP	74		48222 > 22 [SYN] Seq=0 Win=29200	
82		312.264		192.168.50.88		192.168.50.1		TCP	54		22 > 48222 [RST, ACK] Seq=1 Ack=1	
101		462.027		192.168.50.88		142.104.64.205		TCP	74		38544 > 22 [SYN] Seq=0 Win=29200	
102		463.024		192.168.50.88		142.104.64.205		TCP	74		[TCP Retransmission] 38544 > 22 [5	
106		472.196		192.168.50.88		142.104.64.205		TCP	74		58842 > 1922 [SYN] Seq=0 Win=29200	
107		473.196		192.168.50.88		142.104.64.205		TCP	74		[TCP Retransmission] 58842 > 1922	
108		475.2		192.168.50.88		142.104.64.205		TCP	74		[TCP Retransmission] 58842 > 1922	
109		479.212		192.168.50.88		142.104.64.205		TCP	74		[TCP Retransmission] 58842 > 1922	

OK Apply Cancel ⓘ

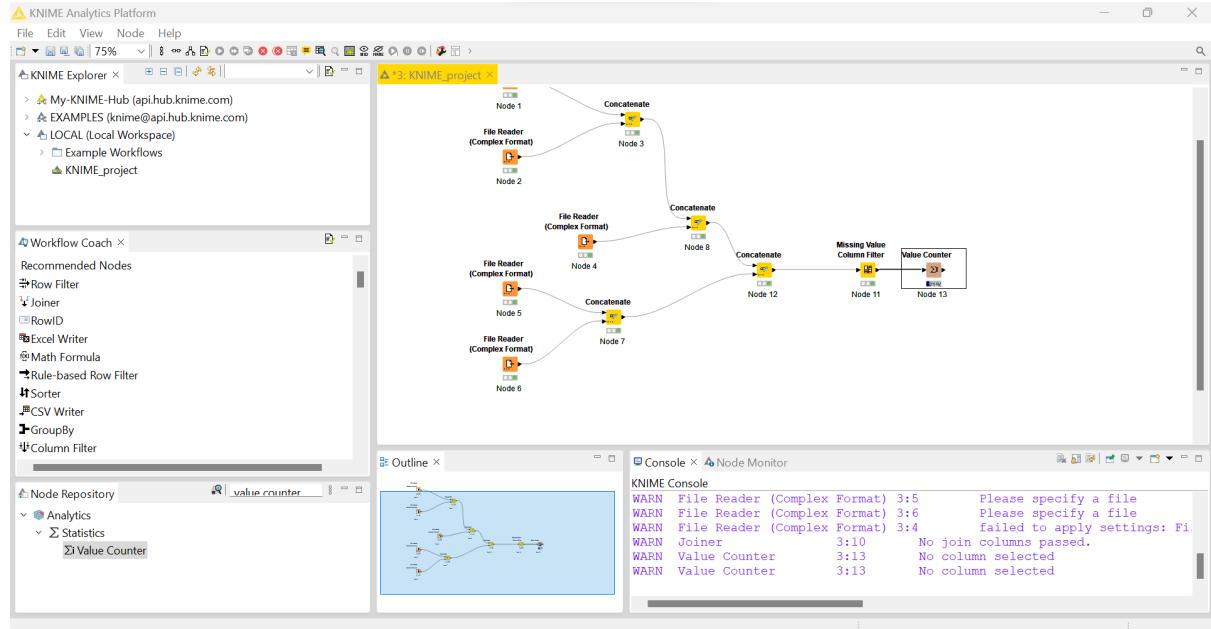
7. Gabungkan kelima data dengan menggunakan concatenate dan data reader



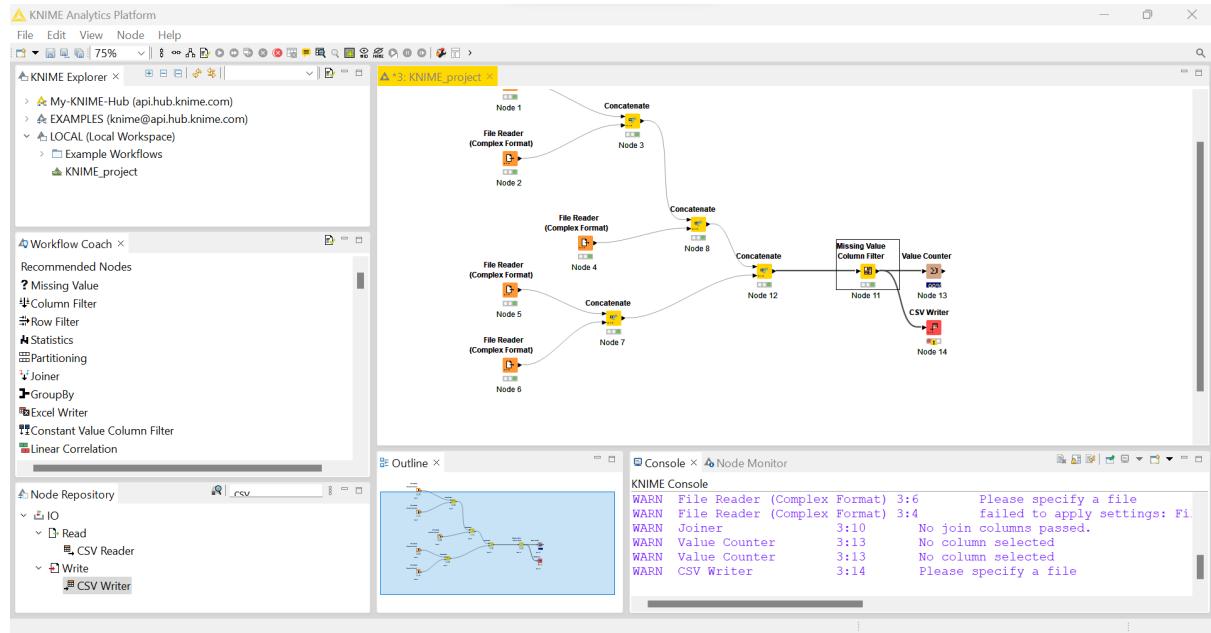
8. Untuk melakukan labeling data normal kita akan menggunakan Node Missing Value. Node ini digunakan untuk mengisi data kosong.



9. Untuk memastikan bahwa kolom label sudah terisi dengan value Malicious atau Normal, dapat menggunakan node Value Counter. Node ini berfungsi untuk menghitung jumlah seluruh value pada kolom terpilih.



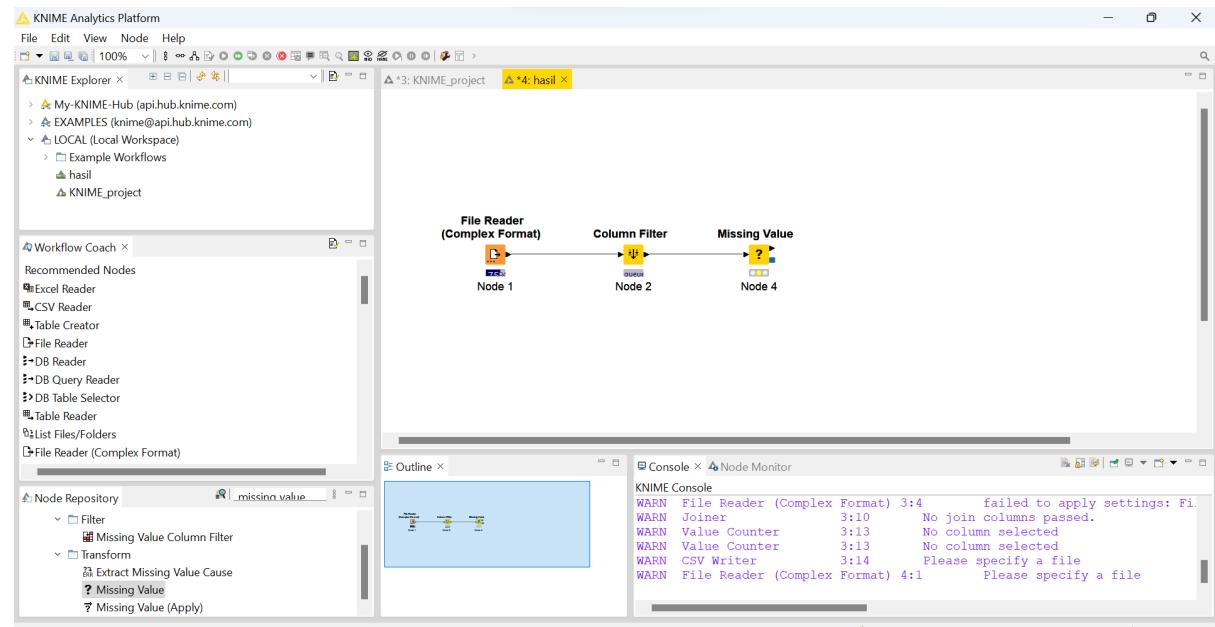
10. Export file ke dalam format .csv dengan menggunakan node CSV Writer



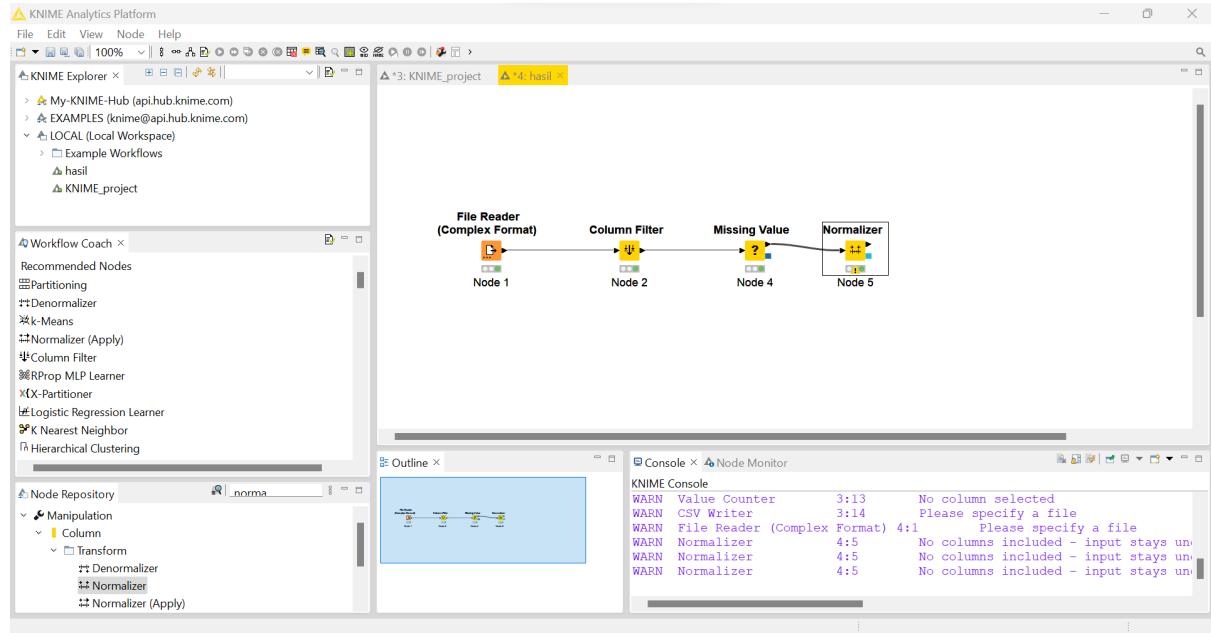
No.	Time	Source	Destination	Protocol	Length	Info
1	"0.000000"	"192.168.50.50"	"192.168.50.88"	"DNS"	"79"	"Standard query 0x5e82 A clients2.google.com"
3	"4.0802921"	"192.168.50.19"	"192.168.50.88"	"DNS"	"81"	"Standard query 0xf4f4 A client-cf.dropbox.com"
9	"1.848537"	"192.168.50.88"	"8.8.4.4"	"DNS"	"97"	"Standard query 0xa0b A updatekeepalive.mcafee.com OPT"
11	"2.096566"	"192.168.50.51"	"192.168.50.88"	"DNS"	"84"	"Standard query 0x4fe5 A www.google-analytics.com"
15	"2.848113"	"192.168.50.88"	"192.168.50.51"	"DNS"	"86"	"Standard query response 0x4234 Server failure A updatekeepalive.mcafee.com"
7	"Time", "Source", "Destination", "Protocol", "Length", "Info"					
8	"309.509587", "192.168.50.1", "192.168.50.88", "TCP", "74", "48221 > 22 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM TStamp=341949681 TSectr=0 WS=64"					
9	"309.509587", "192.168.50.1", "192.168.50.88", "TCP", "54", "22 > 48221 [RST, ACK] Seq=1 Ack=1 Win=0 Len=0"					
10	"312.263619", "192.168.50.1", "192.168.50.88", "TCP", "74", "48222 > 22 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM TStamp=341950369 TSectr=0 WS=64"					
11	"312.263647", "192.168.50.1", "192.168.50.88", "TCP", "54", "22 > 48222 [RST, ACK] Seq=1 Ack=1 Win=0 Len=0"					
12	"462.026672", "192.168.50.88", "142.104.64.205", "TCP", "74", "38544 > 22 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM TStamp=195132807 TSectr=0 WS=128"					
13	"463.024215", "192.168.50.88", "142.104.64.205", "TCP", "74", "[TCP Retransmission] 38544 > 22 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM TStamp=195133057 TSectr=0 WS=128"					
14	"472.199177", "192.168.50.88", "142.104.64.205", "TCP", "74", "58842 > 1922 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM TStamp=195135350 TSectr=0 WS=128"					
15	"473.196245", "192.168.50.88", "142.104.64.205", "TCP", "74", "[TCP Retransmission] 58842 > 1922 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM TStamp=195135600 TSectr=0 WS=128"					
16	"475.200246", "192.168.50.88", "142.104.64.205", "TCP", "74", "[TCP Retransmission] 58842 > 1922 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM TStamp=195136101 TSectr=0 WS=128"					
17	"479.212237", "192.168.50.88", "142.104.64.205", "TCP", "74", "[TCP Retransmission] 58842 > 1922 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM TStamp=195137104 TSectr=0 WS=128"					
18	"487.228249", "192.168.50.88", "142.104.64.205", "TCP", "74", "[TCP Retransmission] 58842 > 1922 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM TStamp=195139108 TSectr=0 WS=128"					

11. Data Pre Processing

Proses dimana data akan dibersihkan (cleaning) karena biasanya didalam suatu data terdapat nilai-nilai yang tidak sempurna atau bahkan terdapat nilai-nilai yang hilang atau kosong yang nantinya akan dapat mempengaruhi proses kedepannya. Pada proses ini kita membutuhkan Node-node berikut : FileReader, Column Filter, Missing Value.

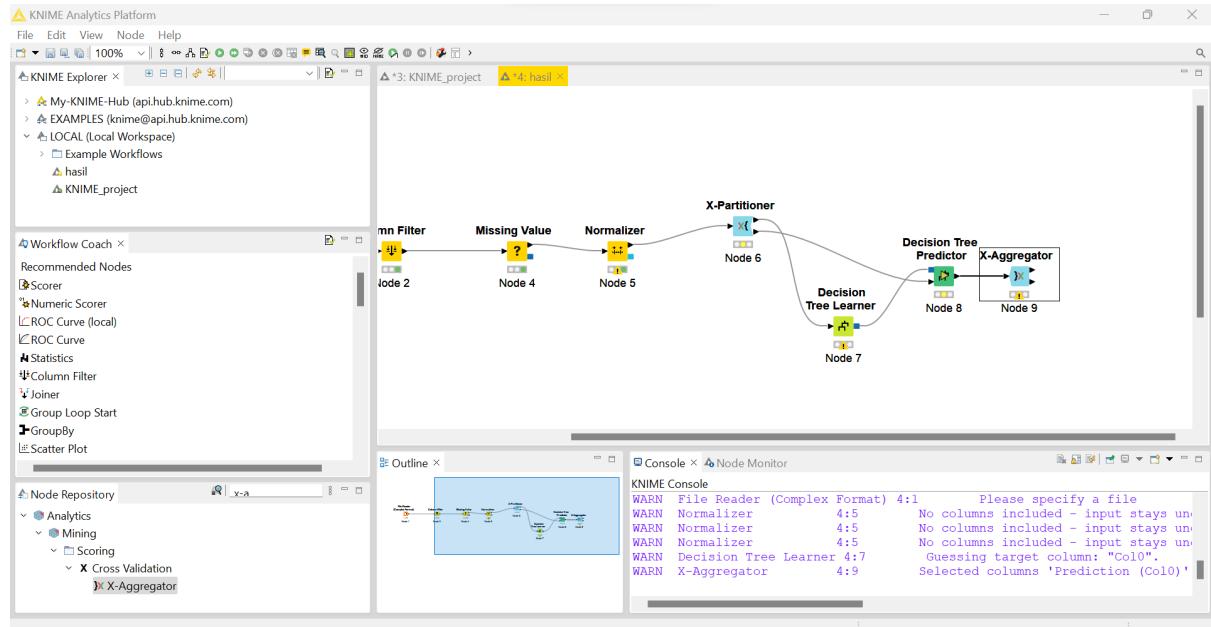


12. Proses data transformation, pada proses ini data akan diubah ke format yang sesuai untuk proses data mining. Node yang digunakan pada tahap ini yaitu Normalizer. Berikut konfigurasinya

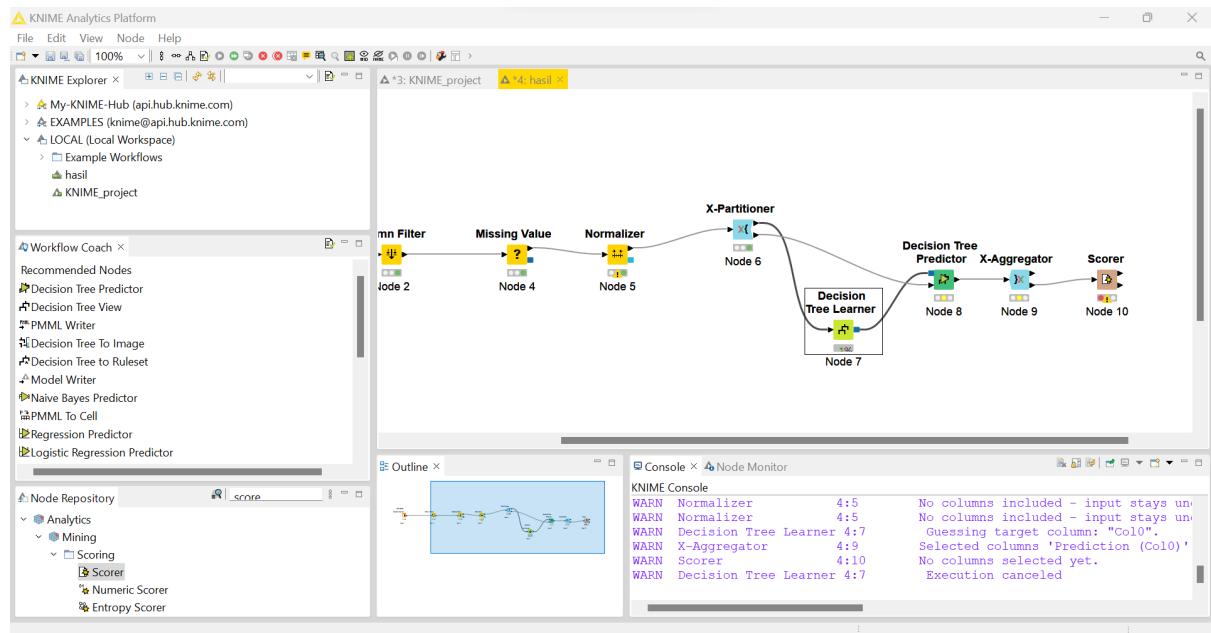


13. Data Mining

Setelah menyelesaikan tahap data transformation, kita akan menjalankan proses Data Mining, dalam proses ini kita akan menggunakan Metode Klasifikasi Decision Tree dengan teknik Cross Validation. Pada proses ini kita membutuhkan Node-node berikut : X-Partitioner, Decision Tree Learner, Decision Tree Predictor, X-Aggregator Sehingga akan membentuk flow seperti ini



14. Node Scorer yang didalamnya terdapat perhitungan untuk melihat seberapa baik model ini dengan menggunakan teknik confusion matrix. Berikut konfigurasinya.



15. Hasil Prediksi

