

Pengisian poin C sampai dengan poin H mengikuti template berikut dan tidak dibatasi jumlah kata atau halaman namun disarankan ringkas mungkin. Dilarang menghapus/memodifikasi template ataupun menghapus

C. HASIL PELAKSANAAN PENELITIAN: Tuliskan secara ringkas hasil pelaksanaan penelitian yang telah dicapai sesuai tahun pelaksanaan penelitian. Penyajian meliputi data, hasil analisis, dan capaian luaran (wajib dan atau tambahan). Seluruh hasil atau capaian yang dilaporkan harus berkaitan dengan tahapan pelaksanaan penelitian sebagaimana direncanakan pada proposal. Penyajian data dapat berupa gambar, tabel, grafik, dan sejenisnya, serta analisis didukung dengan sumber pustaka primer yang relevan dan terkini.

penjelasan di setiap poin.

Enam tahapan siklus hidup pengembangan *data mining* berdasarkan CRISP-DM (*CRoss-Industry Standard Process for Data Mining*):

1. Fase penambahan Tujuan (*Business Understanding Phase*)

Pada tahapan ini penelitian bertujuan untuk mengidentifikasi kebutuhan secara detail yaitu dengan mengidentifikasi pola-pola dataset penerimaan calon mahasiswa baru terdahulu berdasarkan variabel-variabel yang sudah dipilih, seperti terlihat pada tabel di bawah ini:

JK	Alamat		Jurusan	...	
P	Desa. Luwo'o, Kec. Telaga Jaya		Bahasa		
W	Desa Pantungo Kec. Telaga Biru		Keperawatan		
P	Desa. Bulota, Kec. Telaga Jaya		IPS		
...	Gelombang	Pilihan 1	Pilihan 2	Lulus
	3	KEP	PSI	N	
	4	KEP	AGB	Y	
	5	PSI	SI	Y	

...	Agama	Jurusan_lulus	Alamat
I	-		Desa Pantungo Kec. Telaga Biru
I	KEP		Desa. Bulota, Kec. Telaga Jaya
I	PSI		Desa. Olimoo'o, Kec. Batudaa Pantai

2. Fase Pemahaman Data (*Data Understanding Phase*)

Dalam pengumpulan data pada penelitian bersumber dari data calon mahasiswa baru Muhammadiyah

Gorontalo dari tahun 2015 hingga 2017 yang didapat dari tim PMB Muhammadiyah Gorontalo yaitu:

a. Peminat

Tahun	Gelombang 1	Gelombang 2	Gelombang 3	Gelombang 4	Gelombang 5	Jumlah
2015	415	79	105	130	319	1048
2016	675	269	250	177	15	1386
2017	239	201	771	259	-	1470

b. Terdaftar

Tahun	Jumlah Peminat	Jumlah Terdaftar
2015	1048	1009
2016	1386	1370
2017	1470	1442

3. Fase Pengolahan Data (*Data Prepatation Phase*)

Pada tahap ini meliputi proses pengolahan data yaitu data Penerimaan Mahasiswa Baru tahun 2015 sampai 2017. Adapun tahap ini mencakup beberapa proses antara lain:

a. Pemilihan Data

Pada proses ini dilakukan seleksi terhadap beberapa atribut yang akan digunakan untuk pemodelan, yaitu, Jenis Kelamin, Sekolah, Jurusan, Gelombang, Pilihan Jurusan (2 Pilihan), catatan, Agama, Jurusan lulus, Alamat dan Pekerjaan Orang tua.

b. Pembersihan Data

Data awal tidak semuanya berisi data yang utuh atau lengkap. Maka perlu dilakukan proses cleaning, dimana data yang kosong tersebut dihilangkan sehingga tersisa data yang siap diolah.

c. Perubahan Data

Pada tahap ini ada beberapa atribut pada data dapat disederhanakan menjadi atribut baru. Kemudian dilakukan pula konversi nilai pada atribut yang diperlukan:

1. Menerjemahkan nilai **Pilihan1**, **pilihan2** dan **jurusan lulus**

Pilhan1/pilihan2/Jurusan lulus	Nilai Baru
01	Agribisnis
02	Akuntansi
03	Administrasi Publik

04	Akuakultur
05	Budidaya Perairan
06	Bidan Pendidikan
07	Ekonomi Pembangunan
08	Geografi
09	Ilmu Hukum
010	Ilmu Komputer
011	Ilmu keolahragaan
012	Kebidanan
013	Keperawatan
014	Ners
015	Pariwisata
016	PGSD
017	Psikologi
018	Peternakan
019	Sastra Arab
020	Sistem Informasi
021	Sastra Inggris

2. Menerjemahkan Nilai **Agama**

Agama	Nilai Baru
Buddha	B
Hindu	H
Islam	I
Kristen	K
Protestan	P

3. Fase Pemodelan (*Modelling*)

Tahapan ini akan digunakan teknik *Data Mining* dengan metode prediksi menggunakan algoritma C45 dan Nearest Neighbor, yang menghasilkan aturan Prediksi dan variabel paling berpengaruh

dalam memprediksi masuknya calon mahasiswa baru, seperti dibawah ini:

a. Algoritma C45

Pada algoritma ini, merupakan algoritma yang digunakan untuk membentuk pohon keputusan yang terdiri dari sekumpulan aturan untuk membagi sejumlah populasi menjadi kecil.

Tabel Data Calon Mahasiswa Baru

No	Jenis Kelamin	Pilihan Prodi	Pendidikan Terakhir	Jurusan Sekolah	Total Nilai Ujian
1	P	PTN	SMK	Agribisnis Ternak Ruminansia	<90
2	P	BDP	SMU	IPA	<90
3	W	KEP	SMU	IPS	<70
4	P	KEP	SMU	IPS	<70
....
4430	P	PSI	SMU	IPA	<80
4431	P	KEP	SMU	IPA	<90

Tabel Perhitungan Node

Atribut		Jumlah	Lulus	T. Lulus	Entropy	Gain
Total		4431	3748	683	0.62011	
Jenis Kelamin	Pria	1228	1058	170	0.58013	0.00053
	Wanita	3203	2690	513	0.63470	
Pilihan Prodi	AGB	86	71	15	0.66771	0.36466
	AK	87	67	20	0.77781	
	PGSD	477	438	39	0.40835	
	SI	179	155	24	0.56852	
	
	SIG	93	74	19	0.73045	
Pendidi	MA	305	259	46	0.61189	0.00019
	SMK	729	607	122	0.65161	

kan Terakhir	SMU	3397	2882	515	0.61384	
Jurusan Sekolah	Administrasi Perkantoran	94	69	25	0.83560	0.10210
	Agama	4	2	2	1	
	Agribisnis Pengolahan Hasil Pertanian	11	8	3	0.84535	
	
	Teknologi Pengolahan Hasil Pertanian	21	4	17	0.70247	
Total Ujian	Usaha Perjalanan Wisata	2	2	0	0	0.44763
	Tidak Ikut Ujian	77	0	0	0	
	<70	631	96	219	0.94316	
	<80	1453	1407	0	0	
	<90	1679	1657	22	0.10073	

Baris **Total** kolom **Entropy** dihitung dengan persamaan berikut:

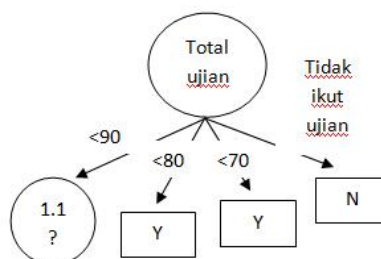
$$Entropy (total) = \left(-\frac{3}{4} \times \log_2\left(\frac{3}{4}\right)\right) + \left(\frac{6}{4} \times \log_2\left(\frac{6}{4}\right)\right) = 0.62011$$

Sementara itu, untuk nilai gain pada jenis kelamin dihitung dengan persamaan berikut:

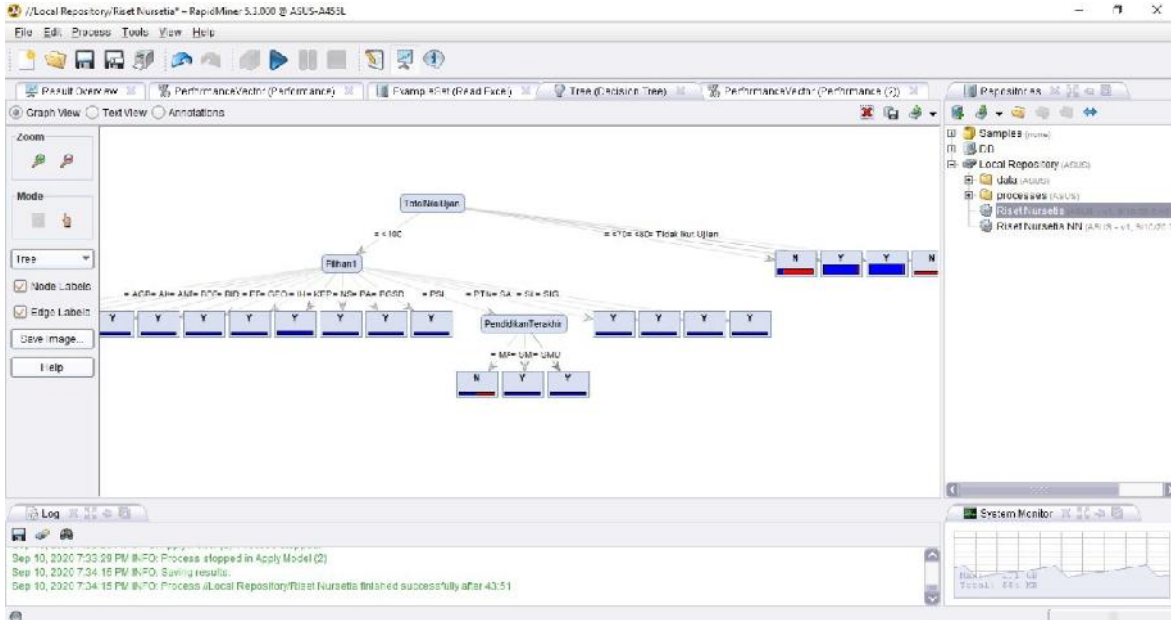
$$Gain(total, jeniskelamin) = Entropy(total) - \sum_{t=1}^n \frac{j_t}{t} \times E(j_t, k_t)$$

$$Gain(total, jeniskelamin) = 0.62011 - \left(\frac{1}{4} \times 0.58013\right) + \left(\frac{3}{4} \times 0.63470\right) = 0.00053$$

Dari hasil perhitungan yang terdapat pada tabel perhitungan node, dapat diketahui bahwa dengan gain tertinggi terdapat pada atribut total ujian yaitu sebesar **0.44763**. Dengan demikian **total ujian** dapat menjadi node akar. Dan pada total ujian terdapat 4 nilai atribut yang dari hasil tersebut dapat digambarkan pohon keputusan sementara seperti berikut:



Hasil pun terlihat sama pada saat kita menguji menggunakan tools Rapidminer yang terlihat pada gambar berikut



b. Algoritma Nearest Neighbor

Pada algoritma ini menggunakan cara pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dan kasus lama, yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada.

Tabel kasus lama

Jenis Kelamin	Pendidikan Terakhir	Agama	Lulus Ujian
P	SMK	Hindu	Y
W	SMU	Islam	N
W	SMU	Islam	Y

Atribut *Lulus ujian* merupakan atribut tujuan.

Bobot antara satu atribut dengan atribut yang lain pada atribut bukan tujuan dapat didefinisikan dengan nilai berbeda.

Tabel Definisi Bobot Atribut

Atribut	Bobot
Jenis Kelamin	0.5
Pendidikan Terakhir	1

Agama	0.75
-------	------

Tabel Kedekatan Nilai Atribut Jenis Kelamin

Nilai 1	Nilai 2	Kedekatan
W	W	1
P	P	1
W	P	0.5
P	W	0.5

Tabel Kedekatan Nilai Atribut Pendidikan Terakhir

Nilai 1	Nilai 2	Kedekatan
SMK	SMK	1
SMU	SMU	1
SMU	SMK	0.4
SMK	SMU	0.4

Tabel Kedekatan Nilai Atribut Agama

Nilai 1	Nilai 2	Kedekatan
Hindu	Hindu	1
Islam	Islam	1
Hindu	Islam	0.75
Islam	Hindu	0.75

Kasus Baru

Jenis Kelamin : P

Pendidikan Terakhir : SMK

Agama : Hindu

Untuk memprediksi apakah calon mahasiswa baru tersebut akan lulus ujian atau tidak, dengan langkah berikut:

1. Menghitung kedekatan kasus baru dengan kasus lama.

Diketahui:

- a. Kedekatan nilai atribut jenis kelamin (Pria dengan Pria) 1
- b. Bobot Atribut Jenis Kelamin 0.5
- c. Kedekatan nilai atribut pendidikan terakhir (SMK dan SMU) 0.4
- d. Bobot atribut pendidikan Terakhir 1
- e. Kedekatan nilai atribut Agama (Hindu dan Hindu) 1
- f. Bobot Atribut Agama 0.75

Dihitung:

$$Jarak = \frac{(a*b)+(c*d)+(e*f)}{b+d+f}$$

$$Jarak = \frac{(1*0.5)+(0.4*1)+(1*0.7)}{0.5+1+0.7}$$

$$Jarak = \frac{1.6}{2.2}$$

$$Jarak = 0.73$$

2. Menghitung kedekatan kasus baru dengan kasus nomor 2

Diketahui:

- a. Kedekatan nilai atribut jenis kelamin (Wanita dengan Pria) 0.5
- b. Bobot Atribut Jenis Kelamin 0.5
- c. Kedekatan nilai atribut pendidikan terakhir (SMK dan SMU) 0.4
- d. Bobot atribut pendidikan Terakhir 1
- e. Kedekatan nilai atribut Agama (Islam dan Islam) 1
- f. Bobot Atribut Agama 0.75

Dihitung:

$$Jarak = \frac{(a*b)+(c*d)+(e*f)}{b+d+f}$$

$$Jarak = \frac{(0.5*0.5)+(0.4*1)+(1*0.7)}{0.5+1+0.7}$$

$$Jarak = \frac{1.4}{2.2}$$

$$Jarak = 0.62$$

3. Menghitung kedekatan kasus baru dengan kasus nomor 3

Diketahui:

- a. Kedekatan nilai atribut jenis kelamin (Wanita dengan Pria) 0.5
- b. Bobot Atribut Jenis Kelamin 0.5
- c. Kedekatan nilai atribut pendidikan terakhir (SMK dan SMU) 0.4
- d. Bobot atribut pendidikan Terakhir 1

- e. Kedekatan nilai atribut Agama (Hindu dan Islam) 0.75
- f. Bobot Atribut Agama 0.75

$$Jarak = \frac{(a*b)+(c*d)+(e*f)}{b+d+f}$$

$$Jarak = \frac{(0.5*0.5)+(0.4*1)+(0.7 *0.7)}{0.5+1+0.7}$$

$$Jarak = \frac{2.1}{2.2}$$

$$Jarak = 0.95$$

- 4. Memilih kasus dengan kedekatan terdekat.

Dari langkah 1,2 dan 3 dapat diketahui bahwa nilai tertinggi adalah kasus pada nomor 3, untuk itu kasus terdekat dari kasus baru adalah kasus 3.

- 5. Menggunakan klasifikasi dari kasus dengan kedekatan terdekat.

Berdasarkan hasil dari langkah 4, maka kasus 3 akan digunakan untuk memprediksi kasus baru dengan kemungkinan calon mahasiswa baru akan *lulus ujian*.

- Berikut Merupakan Hasil Algoritma Nearest Neighbor Menggunakan data Real Menggunakan Excel

Definisi Bobot

Atribut	Bobot
Jenis Kelamin	0.5
Agama	0.75
Pendidikan	1

Kedekatan Nilai Atribut kelamin

Nilai 1	Nilai 2	Kedekatan
P	P	1
W	W	1
P	W	0.5
W	P	0.5

Kedekatan Nilai Atribut Agama

Nilai 1	Nilai 2	Bobot
Hindu	Kristen	0.75
Islam	Hindu	0.75
Kristen	Islam	0.75
Hindu	Hindu	1
Islam	Islam	1
Kristen	Kristen	1

Kedekatan Nilai Atribut Pendidikan

NILAI 1	Nilai 2	Bobot
MA	SMK	0.4
SMU	MA	0.4
SMK	SMU	0.4
MA	MA	1

SMU	SMU	1
SMK	SMK	1

Tabel Data Set

Kelamin	Agama	Pilihan1	Pilihan2	PendidikanTerakhir	NilaiSekolah	LulusUjian
W	I	ANE	SI	SMU	22,13	Y
W	I	BID	EP	SMU	39,20	Y
W	I	KEP	SIG	SMU	6,9	Y
....
P	I	SI	PTN	MA	49.5	Y
W	I	KEP	AK	MA	27,75	Y
P	K	KEP	IH	SMU	6,75	Y
P	I	KEP	IH	SMK	52	N
P	I	KEP	AK	SMU	59.87	Y
P	I	SA	SA	SMU	59,00	Y
W	I	SA	SI	MA	67,6	Y
W	I	PGSD	GEO	SMU	44,50	Y
W	I	PGSD	GEO	SMU	31.12	Y
P	I	KEP	ANE	SMK	49.32	Y
P	K	KEP	PSI	SMU	-	Y
P	I	SI	AGB	SMU	26.75	Y
W	I	KEP	AK	MA	160	Y
P	I	KEP	SI	SMU	36.87	Y
P	I	KEP	SI	SMU	36.87	N

Kasus Baru

Kelamin	Agama	PendidikanTerakhir
W	I	SMU

Tabel Kasus

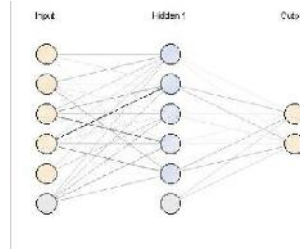
Kelamin	Agama	PendidikanTerakhir
W	I	SMU
W	I	SMU
W	I	SMU
....
P	I	MA
W	I	MA

P	KR	SMU
P	I	SMK
P	I	SMU
P	I	SMU
W	I	MA
W	I	SMU
W	I	SMU
P	I	SMK
P	K	SMU
P	I	SMU
W	I	MA
P	I	SMU
P	I	SMU

No	Kedekatan Nilai Atribut Kelamin	Bobot Atribut Kelamin	Kedekatan Nilai Atribut Agama	Bobot Atribut Agama	Kedekatan Nilai Atribut Pendidikan	Bobot Atribut Pendidikan	Rumus Jarak
1	1	0.5	1	0.75	1	1	1
2	1	0.5	1	0.75	1	1	1
3	1	0.5	1	0.75	1	1	1
4	1	0.5	1	0.75	1	1	1
5	0.5	0.5	1	0.75	1	1	0.888888889
6	1	0.5	1	0.75	0.4	1	0.733333333
...
333	0.5	0.5	1		0.4	1	0.433333333
334	1	0.5	1		1	1	1
335	0.5	0.5	0.75		0.4	1	0.433333333
336	1	0.5	1		0.4	1	0.6
337	0.5	0.5	0.75		0.4	1	0.433333333
338	0.5	0.5	1		0.4	1	0.433333333
339	0.5	0.5	1		0.4	1	0.433333333
340	1	0.5	1		0.4	1	0.6
341	0.5	0.5	0.75		1	1	0.833333333
342	0.5	0.5	1		0.4	1	0.433333333
343	0.5	0.5	1		1	1	0.833333333
344	0.5	0.5	1		1	1	0.833333333
345	1	0.5	1		0.4	1	0.6
346	1	0.5	1		1	1	1
347	1	0.5	1		1	1	1
348	0.5	0.5	1		0.4	1	0.433333333
349	0.5	0.5	0.75		1	1	0.833333333
350	0.5	0.5	1		1	1	0.833333333
351	1	0.5			0.4	1	0.6

➤ Berikut Merupakan Hasil Menggunakan Tool RapidMiner

accuracy: 86.02% +/- 0.26% (mikro: 85.02%)			
	true 0.0	true 1.0	class precision
pred. 0.0	1894	307	86.05%
pred. 1.0	1	1	50.00%
class recall	99.95%	0.32%	



4. Fase Evaluasi (*Evaluation*)

Pada fase ini terlihat hasil dari kedua metode dengan menggunakan Tools RapidMiner

a. Algoritma C4.5

Terlihat pada gambar dibawah ini, Akurasi yang didapatkan sebesar 96.47% dengan Prediksi Yes (Y) true yes 4743 dan True no 76 dengan class Precision 98.42% dan Prediksi No (N) yes 122 dan True no 670 dengan calass Precision 84.60%

accuracy: 96.47% +/- 0.81% (mikro: 96.47%)			
	true Y	true N	class precision
pred. Y	4743	76	98.42%
pred. N	122	670	84.60%
class recall	97.99%	89.81%	

b. Algoritma Nearest Neighbor

Pada gambar dibawah terlihat nilai akurasi dari Algoritma Nearest Neighbor sebesar 86.02% dengan class precision 86.05% dan class precision 50.00%

accuracy: 86.02% +/- 0.26% (mikro: 86.02%)			
	true 0.0	true 1.0	class precision
pred. 0.0	1894	307	86.05%
pred. 1.0	1	1	50.00%
class recall	99.95%	0.32%	

Dari kedua gambar di atas sangat terlihat jelas bahwa akurasi tertinggi terdapat pada algoritma C4.5 ini dikarenakan menurut analisa peneliti algoritma C4.5 ini dalam pemrosesan perhitungan melalui tools RapidMiner terutama, tidak memerlukan proses pengubahan data asli yaitu dari data yang berisi Huruf dan angka, sedangkan untuk algoritma Nearest Neighbor proses pengubahan data itu sangat diperlukan karena saat memproses algoritma tersebut harus bersifat angka dan alasan lainnya yaitu untuk algoritma Nearest Neighbor lebih banyak digunakan untuk proses klasifikasi menurut kedekatan nilai.

5. Fase Penyebaran (*Deployment*)

Pada proses Terakhir ini Peneliti melakukan kesimpulan bahwa terlihat jelas perbandingan antara kedua algoritma tersebut memiliki unggulan pada algoritma C4.5 yang pada pohon keputusannya node teratas ada pada atribut Total ujian, sedangkan terlihat pada data real saat peneliti dan pembantu peneliti melakukan pengumpulan data sangatlah berbeda jauh dengan apa yang diharapkan, yaitu banyaknya data yang kosong atau data yang redundansi saat mahasiswa baru mengisi data-data mereka, terutama pada atribut asal jurusan sewaktu sekolah dan nilai NEM-nya yang itu merupakan atribut penting dalam menentukan hasil terbaik.

D. STATUS LUARAN: Tuliskan jenis, identitas dan status ketercapaian setiap luaran wajib dan luaran tambahan (jika ada) yang dijanjikan. Jenis luaran dapat berupa publikasi, perolehan kekayaan intelektual, hasil pengujian atau luaran lainnya yang telah dijanjikan pada proposal. Uraian status luaran harus didukung dengan bukti kemajuan ketercapaian luaran sesuai dengan luaran yang dijanjikan. Lengkapi isian jenis luaran yang dijanjikan serta unggah bukti dokumen ketercapaian luaran wajib dan luaran tambahan melalui Simlitabmas.

Pada status luaran wajib peneliti menjanjikan publikasi berupa jurnal internasional tidak terakreditasi, yang untuk sementara diselesaikan.

E. PERAN MITRA: Tuliskan realisasi kerjasama dan kontribusi Mitra baik *in-kind* maupun *in-cash* (untuk Penelitian Terapan, Penelitian Pengembangan, PTUPT, PPUPT serta KRUPPT). Bukti pendukung realisasi kerjasama dan realisasi kontribusi mitra dilaporkan sesuai dengan kondisi yang sebenarnya. Bukti dokumen realisasi kerjasama dengan Mitra diunggah melalui Simlitabmas.

.....

.....

.....

.....

.....

F. KENDALA PELAKSANAAN PENELITIAN: Tuliskan kesulitan atau hambatan yang dihadapi selama melakukan penelitian dan mencapai luaran yang dijanjikan, termasuk penjelasan jika pelaksanaan penelitian dan luaran penelitian tidak sesuai dengan yang direncanakan atau dijanjikan.

Kendala yang dihadapi oleh peneliti berupa, penemuan data yang ternyata tidak sesuai dengan yang diinginkan oleh peneliti, yaitu pengisian data calon mahasiswa pada form pendaftaran melalui aplikasi dan lembaran kertas yang sebagian besar calon mahasiswa tidak mengisi nilai akhir mereka saat sekolah, dan asal jurusan saat sekolah, selain itu ada beberapa data yang sudah tercecer atau tidak memiliki backup pada bagian kemahasiswaan

untuk dijadikan sumber data untuk diolah yang itu mengakibatkan dalam pengolahan data peneliti sedikit mendapatkan kesulitan yang mengakibatkan laporan penelitian sedikit terhambat dan pengupload jurnal internasional yang belum bisa dipublis.

G. RENCANA TAHAPAN SELANJUTNYA: Tuliskan dan uraikan rencana penelitian di tahun berikutnya berdasarkan indikator luaran yang telah dicapai, rencana realisasi luaran wajib yang dijanjikan dan tambahan (jika ada) di tahun berikutnya serta *roadmap* penelitian keseluruhan. Pada bagian ini diperbolehkan untuk melengkapi penjelasan dari setiap tahapan dalam metoda yang akan direncanakan termasuk jadwal berkaitan dengan strategi untuk mencapai luaran seperti yang telah dijanjikan dalam proposal. Jika diperlukan, penjelasan dapat juga dilengkapi dengan gambar, tabel, diagram, serta pustaka yang relevan. Jika laporan kemajuan merupakan laporan pelaksanaan tahun terakhir, pada bagian ini dapat dituliskan rencana penyelesaian target yang belum tercapai.

Target yang belum tercapai pada penelitian ini adalah mengupload jurnal seperti dijanjikan, namun jika laporan telah rampug akan di upload sesegera umungkin.

H. DAFTAR PUSTAKA: Penyusunan Daftar Pustaka berdasarkan sistem nomor sesuai dengan urutan pengutipan. Hanya pustaka yang disitasi pada laporan kemajuan yang dicantumkan dalam Daftar Pustaka.

- [1] Retnosari, P. and Jananto, A. (2013) „Implementasi Data Mining Untuk Menemukan Hubungan Antara Kota Kelahiran Mahasiswa Dengan Tingkat Kelulusan Mahasiswa Pada Fakultas Teknologi Informasi Unisbank“, Jurnal Dinamika Informatika, 5(2), pp. 112–121.
- [2] Hartatik. 2015. Penerapan Algoritma *Learning Vector Quantization* untuk prediksi nilai akademis Menggunakan Instrumen AMS (*Academic Motivasion Scale*). Yogyakarta. Teknik Informatika STIMIKAMIKOM Yogyakarta
- [3] Meoritini, Veronika S. 2002, *Data Mining* Data Mining sebagai solusi bisnis, Bandung. Jurnal Integral Vol. 7 No.1 Universitas Katolik Parahyangan.
- [4] Kusri dan Ema Taufiq Lutfi, 2009 Algoritma Data Mining. Universitas Amikom Yogyakarta, Penerbit Andi, Yogyakarta.
- [5] Berry dan Linoff dalam Kusri dan Ema Taufiq Lutfi (2004), 2009 Algoritma Data Mining. Universitas Amikom Yogyakarta, Penerbit Andi, Yogyakarta
- [6] Larose dalam Kusri dan Ema Taufiq Lutfi (2005), 2009 Algoritma Data Mining. Universitas Amikom Yogyakarta, Penerbit Andi, Yogyakarta.