



IMPLEMENTASI DATA MINING TERHADAP PRESTASI BELAJAR MAHASISWA BERDASARKAN PERGAULAN DAN SOSIAL EKONOMI DENGAN ALGORITMA K-MEANS CLUSTERING

Suliman

suliman170892@gmail.com

STIMIK Bina Bangsa, Kendari, Indonesia

Abstrak

Perkembangan teknologi informasi saat ini memungkinkan data dapat disimpan dalam bentuk digital. Pada Perguruan Tinggi jumlah mahasiswa akan terus bertambah seiring berjalannya waktu yang menyebabkan data yang tersimpan terutama data yang berhubungan dengan akademik mahasiswa akan terus bertambah. Data mining merupakan kegiatan yang bertujuan untuk menggali informasi dari tumpukan data yang besar. Data mining merupakan proses analisis data menggunakan perangkat lunak untuk menemukan pola atau aturan tertentu dari sejumlah data dalam jumlah besar yang diharapkan dapat menemukan pengetahuan guna mendukung keputusan. Teknik data mining yang digunakan dalam penelitian ini adalah teknik clustering dengan Algoritma K-Means Clustering dalam proses pengklasifikasian data yang berhubungan dengan akademik mahasiswa untuk dapat mengetahui prestasi dan minat belajar dari mahasiswa yang ada dalam perguruan tinggi. Penelitian ini menghasilkan sebuah model yang dapat digunakan untuk melakukan pengelompokan mahasiswa berdasarkan minat belajarnya sehingga dapat terdeteksi mahasiswa yang bergaul dengan teman yang lebih mengutamakan belajar dan juga dengan status ekonomi yang baik ternyata dapat meningkatkan prestasi belajar mahasiswa.

Kata Kunci: Teknologi Informasi, Data mining, K-Means Clustering.

Abstract

The development of information technology nowadays allows data to be stored in digital form. In Higher Education the number of students will continue to increase over time which causes the data stored, especially data related to student academics, will continue to grow. Data mining is an activity that aims to extract information from large piles of data. Data mining is the process of analyzing data using software to find certain patterns or rules from large amounts of data that are expected to find knowledge to support decisions. The data mining technique used in this research is the clustering technique with the K-Means Clustering Algorithm in the process of classifying data related to student academics in order to be able to find out the achievement and interest in learning of students in college. This study produces a model that can be used to group students based on their learning interests so that students who hang out with friends who prioritize learning and with good economic status can actually improve student achievement.

Keywords: Information Technology, Data mining, K-Means Clustering.

1. Pendahuluan

Perkembangan teknologi informasi terutama yang berhubungan dengan media penyimpanan saat ini berkembang begitu pesat. Media penyimpanan data dapat digunakan oleh sebuah perguruan tinggi untuk menyimpan data yang berhubungan dengan akademik mahasiswa. Data bagi suatu perguruan tinggi sangat bermanfaat jika diolah lebih lanjut untuk menghasilkan sebuah informasi yang dapat digunakan dalam pengambilan keputusan bagi pucuk pimpinan perguruan tinggi tersebut.

Sektor pendidikan harus dijadikan prioritas utama dalam sebuah kebijakan suatu negara. Pendidikan tidak hanya memberikan pengertian akan sesuatu, tetapi juga memberikan pengembangan kepribadian, menanamkan nilai moral, pengetahuan dan keahlian. Seiring dengan berjalannya waktu, data dalam sebuah perguruan tinggi akan terus bertambah sehingga akan menjadi tumpukan data yang menggunung. Jika tidak dikelola dengan baik maka data tersebut hanya berupa data yang tidak ada manfaatnya. Data merupakan hal yang sangat penting dalam melakukan penelitian karena dalam sebuah penelitian selalu melibatkan pengolahan dan analisa data.

Data merupakan sumber utama yang dapat dimanfaatkan untuk menghasilkan informasi. Informasi begitu penting bagi sebuah perguruan tinggi untuk dapat tetap bertahan dalam menghasilkan lulusan yang terbaik. Pengolahan data yang maksimal maka akan menghasilkan informasi yang berkualitas. Data mining memungkinkan pemakai menemukan pengetahuan dalam data database yang tidak mungkin diketahui keberadaannya oleh pemakai. Teknik data mining digunakan untuk memeriksa basis data berukuran besar sebagai cara untuk menemukan pola yang baru dan berguna. Tidak semua pekerjaan pencarian informasi dinyatakan sebagai data mining.

Berdasarkan fenomena inilah maka akan dilakukan sebuah penelitian yang bertujuan untuk membuat sebuah model Data mining yang dapat digunakan untuk melakukan pengelompokan mahasiswa berdasarkan pergaulan dan sosial ekonomi dalam peningkatan prestasi belajar mahasiswa. Data yang digunakan dalam penelitian ini adalah data Mahasiswa Program Studi Sistem Informasi dan Program Studi Sistem Komputer STMIK Bina Bangsa Kendari angkatan 2017 dan Teknik Data mining yang digunakan yaitu Clustering.

2. Metode

2.1 Metode Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini adalah pengamatan (observasi), wawancara (interview) dan studi pustaka.

2.2 Karakteristik Data Warehouse

Karakteristik data warehouse terdiri dari:

1) Berorientasi Subjek

Data warehouse berorientasi pada subjek dan biasanya subjek tersebut meliputi: pelanggan, produk, transaksi atau kegiatan, kebijakan, klaim dan pengguna. Setiap subjek utama secara fisik diimplementasikan sebagai serangkaian tabel yang terkait di data warehouse.

2) Integrasi

Dari semua aspek data warehouse, integrasi merupakan karakteristik yang paling penting. Dalam integrasi, data disuplai dari beberapa sumber yang berbeda ke dalam data warehouse. Pada integrasi data pada sumber diubah, diformat ulang, dirangkum dan sebagainya sesuai dengan kebutuhan data warehouse.

Data yang dimasukkan ke dalam data *warehouse* adalah data yang memiliki konsistensi. Sehingga data pada tingkat aplikasi atau data sumber yang tidak memiliki konsistensi akan dibatalkan. Contohnya adalah penyimpanan data pada jenis kelamin. Jenis data pada jenis kelamin adalah pria dan wanita. Akan tetapi pada prakteknya format data dapat berupa p/l atau 1/0 atau p/w dan lain sebagainya. Perbedaan format penamaan pada jenis kelamin tersebut harus dikonversi menjadi sebuah format yang sama. Sehingga data yang dipindahkan ke dalam data warehouse konsisten. Konsistensi data berlaku untuk semua, seperti konversi penamaan, struktur kunci, pengukuran atribut, dan karakteristik fisik data.

3) Non Volatile

Data diperbaharui atau diubah dalam lingkungan operasional merupakan hal yang biasa. Namun data pada data warehouse memiliki karakteristik yang berbeda. Data warehouse dapat dimuat dan diakses tetapi tidak dapat diperbaharui. Maksudnya adalah data tidak dapat diubah, dihapus atau ditambah.

4) Time Variant

Time Variant atau rentang waktu menyiratkan bahwa setiap unit data dalam data warehouse akurat pada beberapa waktu. Dalam beberapa kasus, record ditandai oleh waktu. Dalam kasus lain, record memiliki tanggal transaksi. Tapi dalam setiap kasus, ada beberapa bentuk penandaan waktu untuk menunjukkan waktu selama record data akurat .

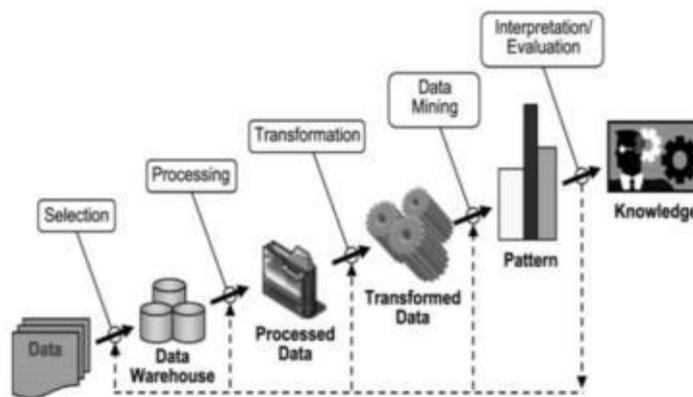
2.3 Data Mining

Menurut [1] data mining (DM) adalah gabungan sejumlah disiplin ilmu computer (ACM2006), yang mendefinisikan sebagai proses penemuan pola-pola baru dari kumpulan-kumpulan data sangat besar, meliputi metode-metode yang merupakan irisan dari *artificial intelligence*, *machine learning*, *statistic* dan *database systems* (ACM 2006). Data mining atau penambahan kata adalah teknik yang relatif cepat dan mudah untuk menemukan pola dan/atau relasi antar data, secara otomatis. Dengan menggabungkan empat disiplin ilmu komputer seperti definisi diatas, pengetahuan bisa ditemukan dalam lima proses berurutan: seleksi, pemrosesan, transformasi, data mining, dan interpretasi/evaluasi.

Dari definisi diatas, maka dapat disimpulkan bahwa secara umum data mining adalah suatu kegiatan analisa data untuk mencari suatu pola tertentu, dengan jumlah data yang besar dan bertujuan untuk menghasilkan informasi yang dapat digunakan dan dikembangkan lebih lanjut.

Data mining merupakan inti dari proses Knowledge Discovery Databases (KDD), meliputi dugaan algoritma yang mengeksplor data, membangun model dan menemukan pola yang belum diketahui. KDD merupakan penyelesaian masalah dengan menganalisa data yang ada pada database dengan data tersimpan secara elektronik dan pencariannya dilakukan otomatis seperti pada computer.

Tahapan pada proses KDD pada database menurut [2] seperti pada Gambar 1 dibawah ini :



Gambar 1. Tahapan proses KDD

1. Data Selection

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining, disimpan suatu berkas, terpisah dari basis data operasional.

2. Pre-processing/Cleaning

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi fokus KDD. Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

3. Transformation

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses coding dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. Data Mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. Interpretation/ Evaluation

Pola informasi yang dihasilkan dari proses data mining, perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya. Langkah terakhir KDD adalah mempresentasikan pengetahuan dalam bentuk yang mudah dipahami pengguna.

2.4 Clustering

Clustering merupakan suatu teknik data mining yang mengelompokkan data ke dalam beberapa kelompok (grup atau cluster atau segmen) yang tiap cluster dapat ditempati beberapa anggota bersama-sama. Pada dasarnya clustering merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (*similarity*) antara satu data dengan data yang lain. Clustering merupakan salah satu metode data mining yang bersifat tanpa arahan (*unsupervised*), maksudnya metode ini diterapkan tanpa adanya latihan (*training*) dan tanpa ada guru (*teacher*) serta tidak memerlukan target output. Dalam data mining ada dua jenis metode clustering yang digunakan dalam pengelompokan data, yaitu *hierarchical clustering* dan *non-hierarchical clustering*.

Hierarchical clustering adalah suatu metode pengelompokan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang memiliki kedekatan kedua. Demikian seterusnya sehingga cluster akan membentuk semacam pohon dimana ada hierarki (tingkatan) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah cluster.

Berbeda dengan metode *hierarchical clustering*, metode *non-hierarchical clustering* justru dimulai dengan menentukan terlebih dahulu jumlah cluster yang diinginkan (dua cluster, tiga cluster, atau lain sebagainya). Setelah jumlah cluster diketahui, barulah proses cluster dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan *K-Means Clustering*. Definisi *K-Means* merupakan salah satu metode data clustering non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih cluster/kelompok. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan cluster/kelompok yang lain sehingga data yang berada dalam satu cluster/kelompok memiliki tingkat variasi yang kecil [3].

2.5 Metode Analisis K-Means

Dalam statistic dan mesin pembelajaran, pengelompokan *K-Means* merupakan metode analisis kelompok yang mengarah pada pemartisian N objek pengamatan ke dalam K kelompok (cluster) dimana setiap objek pengamatan dimiliki oleh sebuah kelompok dengan mean (rata-rata) terdekat, mirip dengan algoritma *Expectation-Maximization* untuk *Gaussian Mixture* dimana keduanya mencoba untuk menemukan pusat dari kelompok dalam data sebanyak literasi perbaikan yang dilakukan oleh kedua algoritma. Pengelompokan data dengan metode *K-Means* ini secara umum dilakukan dengan langkah berikut [4].

- 1) Tentukan jumlah kelompok.
- 2) Alokasikan data ke dalam kelompok secara acak.
- 3) Hitung pusat kelompok (sentroid/rata-rata) dari data yang ada di masing-masing kelompok.
- 4) Alokasikan masing-masing data ke sentroid/rata-rata tersebut
- 5) Kembali ke langkah 3, jika masih ada data yang berpindah kelompok, atau ada perubahan nilai sentroid diatas nilai ambang yang ditentukan, atau apabila perubahan nilai pada fungsi objektif yang digunakan masih diatas nilai ambang yang ditentukan.

2.6 Pergaulan

Pergaulan memiliki pengaruh dalam membentuk suatu komponen kepribadian seseorang, pergaulan yang dilakukan oleh seseorang tanpa disadari akan membentuk dirinya, menyerupai lingkungan

bergaulnya. Jadi kepribadian seseorang dapat dilihat dari pergaulannya, apakah pergaulan yang positif atau pergaulan yang negatif [5]. Pergaulan memiliki ragam pola yang terarah serta ragam pola yang tidak terarah.

Ragam pola terarah merupakan pergaulan yang memiliki pola positif yaitu kearah lingkungan yang baik serta tidak melanggar peraturan yang berlaku di masyarakat dan membawa pengaruh positif. Misalnya kerjasama atau gotong royong membersihkan lingkungan atau membentuk komunitas gemar membaca. Sedangkan ragam pola tidak terarah adalah pergaulan yang memiliki pola negatif biasanya identik dengan sesuatu yang bebas dan tanpa aturan. Biasanya pergaulan ini sering dilakukan oleh remaja dan melanggar peraturan yang ada di masyarakat.

2.7 Faktor Sosial Ekonomi

Sosial ekonomi adalah segala sesuatu hal yang berhubungan dengan tindakan pemenuhan kebutuhan masyarakat, keadaan sosial ekonomi orang berbeda-beda. Ada yang memiliki sosial ekonomi yang rendah dan ada pula yang memiliki sosial ekonomi yang tinggi. Keadaan sosial ekonomi dapat juga dikatakan sebagai suatu keadaan atau kedudukan yang di atur secara sosial dan menetapkan seseorang di suatu posisi tertentu, misalnya pemberian pekerjaan .

Terdapat beberapa variabel yang sering digunakan sebagai indikator untuk mengukur status sosial ekonomi antara lain pekerjaan, pendapatan, dan tingkat pendidikan [6]. Ukuran atau kriteria yang dipakai untuk menggolongkan anggota masyarakat ke dalam kelas tertentu adalah kekayaan, kekuasaan/jabatan, kehormatan, dan pendidikan/ ilmu pengetahuan.

Pada status sosial ada beberapa kriteria yang biasa dipakai untuk menggolongkan anggota masyarakat/keluarga ke dalam suatu lapisan [7].

1. Ukuran kekayaan, Lapisan masyarakat teratas merupakan yang memiliki kekayaan paling banyak.
2. Ukuran kekuasaan, Barang siapa memiliki kekuasaan atau memiliki wewenang terbesar menempati lapisan atas.
3. Ukuran kehormatan, Orang yang paling disegani memiliki tempat teratas, ukuran ini terlepas dari jumlah kekayaan dan kekuasaan. Hal ini banyak dijumpai dalam masyarakat tradisional. Biasanya adalah orang tua, dan yang pernah berjasa.

3. Hasil Dan Pembahasan

Untuk melakukan data mining terhadap data maka dibutuhkan beberapa tabel. Tabel ini dalam database sebenarnya sudah ada, namun tabel yang dibuat ini lebih kecil, karena hanya terdiri dari field-field yang diperlukan untuk proses *Data mining* saja. Rancangan terhadap database yang akan digunakan untuk proses dijelaskan di bawah ini.

3.1 Rancangan Database

Database yang dirancang dalam penelitian ini terdiri dari beberapa tabel. Tabel-tabel dalam rancangan dapat diuraikan sebagai berikut:

1) Tabel Mahasiswa (tb_mhs)

Tabel ini digunakan untuk menampung data-data pribadi mahasiswa. Adapun atribut (*field*) dapat diuraikan sebagai berikut:

Tabel 1. Tabel Mahasiswa

Nama	Type	Width
nim*	varchar	9
nama	varchar	40
jenis_kelamin	varchar	2
temp_tgl_lahir	date	50
kd_prodi	varchar	3
alamat	varchar	250
asal_sek	varchar	20
Penghasilan_ortu	varchar	20

2) Tabel hasil nilai (tb_nilai)

Tabel 2. Tabel hasil nilai

Nama	Type	Width
Nim*	varchar	9
tugas	int	3
kehadiran	Int	6
latihan	int	6
Uts	int	6
uas	int	6
nilai_angka	int	6
nilai_huruf	varchar	1

3.2 Seleksi Data (Data Selection)

Sumber data yang digunakan dalam penelitian berasal dari data mahasiswa angkatan 2017 pada Program Studi Sistem Informasi dan Program Studi Sistem Komputer. Beberapa jenis data diperoleh dari sistem yang berjalan yang digunakan untuk penelitian, indikator penentu dalam pengelompokan data keluaran yang diinginkan.

1) Data Mahasiswa (tb_mhs)

A	B	C	D	E	F	G	H	I
NO	NAMA MAHASISWA	STAMBUK	JK	TEMPAT DAN TGL LAHIR	PRODI	ALAMAT	ASAL SEKOLAH	PENGHASILAN ORANG TUA
1	WA ODE NURHAJIA	201751083	P	LAHIMBUA, 10 SEPTEMBER 1998	SI	KEL. WATULONDO	SMA N 4 KONSEL	3000000
2	NASRI	201751017	L	TOBURU, 09 JUNI 2000	SI	ANDUONOHU	SMA N 1 KONSEL	2000000
3	EVI PURNAMA SARI	201751085	P	WAWATU, 05 DESEMBER 1990	SI	MORAMO	SMK N 4 KONSEL	1000000
4	SARTIKA	201751044	P	NOKO, 13 NOVEMBER 1996	SI	JL.H.E.A MOKODOMPIT, LRG PELANGI	SMA N 2 KENDARI	500000
5	HEPIN	201751075	L	TOBIMEITA, 17 MEI 1994	SI	ANGGALOMELAI	SMA N 11 KONSEL	1000000
6	YUSRI YANSA	201751008	L	KENDARI, 23 OKTOBER 1985	SI	KENDARI	SMA N 1 ANGKASA	1000000
7	ERLITA	201751076	P	RAHA, 28 FEBRUARI 1996	SI	JLN.MADE SABARA	MAN 1 KENDARI	1000000
8	MUHAMMAD FEBRIAN MUNIR	201751002	L	LAFEU, 10 OKTOBER 1997	SI	LAFEU	SMA N 1 UNAAHA	3000000
9	SALHUDIN HAULIN SANGGA	201751021	L	ASAO, 17 FEBRUARI 1999	SI	ASAO	MAN 1 KENDARI	3000000
10	LISMA HIDAYAH HALIM	201751014	P	ASAO, 14 NOVEMBER 1998	SI	ASAO	SMKS EKA BHAKTI KENDARI	4000000
11	NOVITA WULANDARI	201751025	P	WAWOLURI, 05 SEPTEMBER 1999	SI	JL.PATIMURA	SMKN 6 KENDARI	1000000
12	NOVAL NIATI	201751019	P	WANGGUDU, 02 JANUARI 1998	SI	KONUT	SMK N 1 WANGI-WANGI	3000000
13	VIVIT ITMIFATIHA	201751013	P	GUNUNG SARI, 05 MEI 1999	SI	JL.MARTANDU, LRG.SAKURA	SMK N 1 WANGI-WANGI	2500000
14	LA ODE NARWI	201751007	L	LAKOMEA, 05 MARET 1998	SI	SAMPARA	SMA N 20 KONAWA SELATAN	500000
15	MUH. DENNY P	201751001	L	KASIPUTE, 26 FEBRUARI 1999	SI	KEC.RUMBIA TENGAH, KAB BO	SMA N 1 KALIDUPA	2500000
16	ARDI IRWANTO	201750995	L	MALUKU, 29 JUNI 1993	SI	LRG.ANGGREK	SMK N 1 UNAAHA	1000000
17	DEVIKA MUTMAINA	201751085	P	LANTAGI, 26 JUNI 1996	SI	LANTAGI	SMA N 3 KENDARI	700000
18	NOKLEU	201751022	P	WAWORAHA, 10 MEI 1998	SI	LASOLO	SMK N 1 KENDARI	700000
19	ARSAD	201751058	L	LALONGGASU, 21 DESEMBER 1999	SI	JL.SULTAN QAIMUDDIN KDI	SMK N 1 KENDARI	1000000
20	AJI PRATOMO	201751094	L	KENDARI, 27 SEPTEMBER 1997	SI	JL.MY JEND KATAMISO, DS.PUO	PKBM MENDIDHOA	500000
21	BUSHRAN MATAPA	201751124	L	LAROENAI, 21 FEBRUARI 2000	SI	WUA-WUA, LRG HIKMA	SMK N 1 BAU-BAU	500000
22	MUH. ABDUH SYAFI'I	201751089	L	PALINGI, 13 JUNI 1996	SI	JL.BERINGIN	SMA N 01 BOMBANA	2500000
23	SADLI SIDIQ	201751001	L	NAMBO, 12 OKTOBER 1997	SI	KEL.PETOAHA, KEC.NAMBO	SMA N 1 LATAMBAGA	1000000
24	WAYS AL QARNI	201751086	L	WATUDEMA, 20 NOVEMBER 1999	SI	KONSEL	SMA N 5 KENDARI	2000000
25	AKBAR	201751016	L	KENDARI, 24 MARET 1993	SI	PUWATU	SMA N 2 PARIGI	2000000
26	RIFKI DG KAENG	201751024	L	DEPEHA, 20 NOVEMBER 1998	SI	DS.PERSIAPAN WONUA MORINI	SMK N 4 KENDARI	1000000
27	LA ODE MUHARDIN	201751025	L	ANDOWIA, 05 SEPTEMBER 1999	SI	ANDOWIA	SMA N 19 KONAWA SELATAN	700000
28	ARDIANSYAH	201751026	P	NOKO, 27 AGUSTUS 1999	SI	KENDARI	SMK N 4 BOMBANA	2000000
29	DEVI NUR TASYA	201751027	P	TONE-TONEA, 07 NOVEMBER 1999	SI	WAWONII	SMAN 1 MALIGANO	1000000

(Sumber: pengolahan data mahasiswa STMIK Bina Bangsa Kendari)

Gambar 2. Data Mahasiswa

2) Data Nilai (tb_nilai)

NO	NAMA MAHASISWA	STAMBUK	UTS						UAS					
			NILAI			JUMLAH	NILAI HURUF	NILAI			JUMLAH	NILAI HURUF		
			TUGAS 10%	LATHAN 15%	UTS 20%			KEHADIRAN 16%	TUGAS 10%	LATHAN 15%			UAS 20%	KEHADIRAN 16%
1	WA ODE NURHAJIA	201751083	10	4	15	16	45	B					90	B
2	NASRI	201751017	7	15	18	16	56	A			50		106	A
3	EVI PURNAMA SARI	201751085	3	4	10	16	33	C			45		78	B
4	SARTIKA	201751044	3	7	18	16	44	B			50		94	B
5	HEPIN	201751075	3	4	17	16	40	B			50		90	B
6	YUSRI YANSA	201751008	2		5	8	15	EROR		5	55		75	B
7	ERLITA	201751076	3		18	12	33	C			45		78	B
8	MUHAMMAD FEBRIAN MUNIR	201751002	3	12	20	12	47	B			50		97	A
9	SALHUDIN HAULIN SANGGA	201751021	3	12	20	12	47	B			45		92	B
10	LISMA HIDAYAH HALIM	201751014	7	15	14	36	C			45		81	B	
11	NOVITA WULANDARI	201751025	7	10	12	29	C		1	45		75	B	
12	NOVAL NIATI	201751019	7	5	8	20	C			50		70	C	
13	VIVIT ITMIFATIHA	201751013	4	15	12	31	C			40		71	C	
14	LA ODE NARWI	201751007	4	5	10	19	EROR					19	EROR	
15	MUH. DENNY P	201751001	4	5	10	19	EROR					19	EROR	
16	ARDI IRWANTO	201750995	4		8	12	EROR					12	EROR	
17	DEVIKA MUTMAINA	201751085	4	10	12	26	C					26	EROR	
18	NOKLEU	201751022	4	5	12	21	C					21	EROR	
19	ARSAD	201751058	4	10	12	26	C			50		76	B	
20	AJI PRATOMO	201751094	12	10	6	28	C			45		73	C	
21	BUSHRAN MATAPA	201751124	15	20	14	49	B		26			75	B	
22	MUH. ABDUH SYAFI'I	201751089	10	10	10	30	C		5	40		75	B	
23	SADLI SIDIQ	201751001	5	8	13	EROR			50	12		75	B	
24	WAYS AL QARNI	201751086	5	2	7	EROR			23	45		75	B	

(Sumber: pengolahan data nilai mahasiswa STMIK Bina Bangsa Kendari)

Gambar 3. Data Nilai

3.3 Praproses Data

3.3.1 Pembersihan Data (*Data Cleaning*)

Tahap kedua pada proses *Data mining* adalah *cleaning data* yaitu melakukan pembersihan data terhadap *noise* yang ditemukan berupa *missing value*, *inkonsisten data*, dan *redundant data*.

ID	NAMA	TGL LAHIR	JENIS KELAMIN	ALAMAT	SEKOLAH	SKOR	
47	LA UMAR	201752026	L	TAOSU, 10 DESEMBER 1998	SI LRG.PELANGI	SMA N 2 PARIGI KAB.MUNA	2000000
48	MUHAMMAD IKBAL	201752071	L	KASIPUTE, 09 NOVEMBER 1999	SI JL.HALUOLEO, NANGA-NANGA	SMA N 3 KULISUSU KAB.BUTUR	2000000
49	ANDRI	201752021	L	PUNGGALUKU, 12 APRIL 1998	SI BTN.MEKAR.ASRI	SMK N 1 KULISUSU KAB.BUTUR	1000000
50	ADAM	201752005	L	LANOWULU, 01 MEI 1999	SI KEL.TAPUAHI, BOMBANA	SMA N 1 KULISUSU KAB.BUTUR	1000000
52	HASNITA	201752018	P	LAWEI, 06 DESEMBER 1999	SI KEL.PUNGGALUKU, KEC.LAEY	SMA N 10 KONAWA SELATAN	3000000
53	ARWAN NADZAR	201752031	L	KENDARI, 28 NOVEMBER 1996	SI DS.TINANGGEEA KONSEL	SMA N 1 KULISUSU UTARA	3000000
54	IRNAWATI	201752084	P	WAWOBENDE, 13 JULI 1995	SI LAWEI	SMK N 7 KONSEL	3000000
55	ELMIATI MUSA	201752008	L	RANOMEETO, 19 OKTOBER 1999	SI KEC.SAMPARA, KAB.KONAWA	PKBM.MENDIDOHA	1000000
56	WARDA YANTI	201752020	P	UNAAHA, 07 SEPTEMBER 1999	SI KEC.LANDONO, KAB.KONSEL	SMK N 7 KONSEL	3000000
57	MINTIN NURJANAH	201752137	P	KENDARI, 07 OKTOBER 1998	SI RANOMEETO	SMK N 1 WAWONII	2000000
58	YULINAR	201752099	P	LALAMBIA, 24 NOVEMBER 1997	SI KENDARI	MA LEMO BAJO	3000000
59	ARNI WANTI MANSUR	201752080	L	LAMOOSO, 20 DESEMBER 1998	SI KONAWA	MA AL-IKHLAS WAWOTOBII	1000000
60	TRI EPRIDA PUTRI	201752083	L	WAHA, 11 DESEMBER 1999	SI JL.PASAENO	SMK TELKOM KENDARI	2000000
61	RYAN FEBRIANTO	201752006	L	WAHA, 23 DESEMBER 1997	SI JL.LALAMBIA	SM.EA	1000000
62	MUHAMMAD AWALUDDIN	201752079	L	UNAAHA, 23 MARET 1999	SI DS.LAMOOSO, KEC.ANGATA, K	SMK N 6 KENDARI	1000000
63	LENI MARLINA	201752128	P	LAGIWAIE, 11 NOVEMBER 1999	SI LR.DAMAI	SMK N 9 KONSEL	3000000
64	RIZAL	201752131	L	PEMALANG, 07 MEI 1997	SI LR.DAMAI	SMA N 1 KAMBOWA	1000000
65	NUR WAHID	201752027	L	SOPANDO, 17 AGUSTUS 1997	SI DS.TANJUNG TIRAM, MORAMO	SMK N 4 KENDARI	3000000
66	ANDI	201752028	L	NOKO, 14 AGUSTUS 1999	SI KALIDUPA	SMA N 1 POLI-POLIA	1000000
67	ROSIDA	201752029	L	KENDARI, 19 MARET 1999	SI AMBEKAI, KEC.UNAAHA, KAI	SMK N 2 BOMBANA	1000000
68	MUH. ARIF. H	201752030	L	ALEBO, 27 MARET 1994	SI PAMATA RAYA	SMA N 1 ASERA	1000000
69	ARIF ANUGRAH	201752031	L	DIMBA, 04 APRIL 1999	SI JL.SALANGGA	SMA N 10 BOMBANA	5000000
70	SRY WAHYUNI	201752032	P	POKURUMBA, 15 FEBRUARI 1998	SI PERUMNAS POASIA BLOK C.225	PAKE C, DPN KOTA KENDARI	3000000
71	MARJAN ADI SAPUTRA	201752033	L	ULUNGGOLAKA, 05 JULI 1999	SI KONSEL	SMK N 1 KENDARI	1000000
72	IRMAWAN	201752034	L	KENDARI, 26 MARET 1999	SI JL.LAWATA	MA INDOTEC KENDARI	1000000
73	BAGAS PURWANTO	201752035	L	RETE, 16 MARET 1999	SI ANDUONOBU	SMK N 3 KENDARI	5000000
74	JUMRAN	201752036	L	TANGGOBU, 07 JUNI 1998	SI KOLAKA	SMA N 1 RAHA	5000000
75	ADE PUTRA	201752037	L	KENDARI, 15 NOVEMBER 1997	SI JL.KAPTEN PIERE TENDEAN	SMA N 3 POSO	5000000
76	MUH.AL.GAZALI	201752038	L	SIKELI, 19 SEPTEMBER 1999	SI ANDUONOBU	SMA N 1 UNAAHA	1000000
77	HAMDI	201752038	L	MALAYSIA, 09 SEPTEMBER 1999	SI POASIA	SMA N 1 UNAAHA	1000000

Gambar 4. Proses cleaning data

Pada gambar di atas merupakan hasil proses data yang sudah di *cleaning*, sehingga sudah tidak ada lagi data yang *missing value*. Hal ini dilakukan karena data yang *missing value* tidak memberikan informasi apapun jika dipertahankan keberadaannya. Tahapan proses *cleaning data* atau disebut juga filter data dari perancangan *tb_mhs*, *tb_khs* dan *tb_mk*. Dari tahap *cleaning data* yang telah dilakukan diatas diperoleh jumlah data pada *tb_mhs* sebanyak 76 record.

3.3.2 Integrasi Data

Tahap ketiga pada proses data mining adalah tahap integrasi data yaitu penggabungan data dengan tujuan memindahkan seluruh data yang telah di *cleaning* menjadi satu tabel. Pada tahap ini dari tiga tabel akan di integrasikan untuk mendapatkan data source mahasiswa dan data source jumlah nilai keseluruhan. Untuk mendapatkan data source mahasiswa proses penggabungan data dilakukan dengan merelasikan table pergaulan, tabel pendapatan orang tua atau mahasiswa dan table nilai prestasi belajar mahasiswa matakuliah sistem operasi. Hasil dari integrasi atau *join* antar tabel diatas maka didapat data source mahasiswa, seperti pada gambar berikut.

NO	NAMA MAHASISWA	STAMBUK	NILAI PERGAULAN TEMAN	NILAI PENDAPATAN ORANG TUA/MAHASISWA	NILAI PRESTASI BELAJAR SISTEM OPERASI
1	WA ODE NURHAJJA	201751083	108	3000000	90
2	NASRI	201751017	111	2000000	106
3	EVI PURNAMA SARI	201751085	112	1000000	78
4	SARTIKA	201751044	110	500000	94
5	HEPIN	201751075	108	1000000	90
6	YUSRI YANSA	201751008	112	1000000	75
7	ERLITA	201751076	115	1000000	78
8	MUHAMMAD FEBRIAN MUNIR	201751002	103	3000000	97
9	SALHUDIN HAULIN SANGGA	201751021	105	3000000	92
10	LISMA HIDAYAH HALIM	201751014	114	4000000	81
11	NOVITA WULANDARI	201751025	106	1000000	75
12	NOVAL NIATI	201751019	117	3000000	70
13	VIVIT ITMIFATIHA	201751013	110	2500000	71
14	LA ODE NARWI	201751007	92	500000	19
15	MUH. DENNY P	201751001	111	2500000	19
16	ARDI IRWANTO	201750995	115	1000000	12
17	DEVIKA MUTMAINA	201751085	112	700000	26
18	NOKLELI	201751022	113	700000	21
19	ARSAD	201751058	113	1000000	76
20	AJI PRATOMO	201751094	111	500000	73

Gambar 5. Integrasi Data

3.3.3 Tranformasi Data

Tahap keempat pada proses Data mining adalah tahap tranformasi data yaitu data diubah menjadi bentuk yang sesuai untuk diproses dalam Data mining. Dalam penelitian ini data yang akan diproses dari database mysql untuk digunakan di tool SSIS maka format tersebut diubah menjadi *mdf*.

3.4 Sosial Ekonomi Keluarga

1) Pendidikan orang tua

Berdasarkan hasil analisis, rata-rata pendidikan orang tua yaitu SD. Rincian tentang pendidikan orang tua dapat dilihat pada tabel berikut ini.

Tabel 3. Pendidikan orang tua

NO	Jenjang pendidikan	Ayah	presentasi	Ibu	Presentasi
1	Tidak sekolah	7	9,2	2	2,6
2	SD	44	57,9	25	32,9
3	SMP	10	13,2	20	26,3
4	SMA	11	14,5	20	26,3
5	Diploma	1	1,3	2	2,6
6	Sarjana	3	3,9	7	9,2
	Total	76	100	76	100

Perbandingan antara latar belakang pendidikan ibu dengan ayah terlihat bahwa latar belakang pendidikan ibu lebih tinggi dari pada pendidikan ayah. Jumlah ibu dari mahasiswa yang sarjana sebanyak 7 orang sedangkan ayah sebanyak 3 orang dan diploma sebanyak 1 orang ayah dan 2 orang ibu. Jenjang pendidikan ayah lebih banyak tamatan SD yaitu 44 orang sedangkan ibu lebih merata yaitu SMP 20 orang dan SMA 20 orang dan SD sebanyak 25 orang.

2) Pekerjaan orang tua

Pekerjaan yang paling banyak adalah petani, baik pekerjaan orang tua perempuan maupun orang tua laki-laki. Sebanyak 37 orang atau 48,7 persen ibu dan 52 orang atau 68,4 persen ayah bekerja sebagai petani. Walaupun daerah Mentawai berupa kepulauan namun dari data yang diperoleh hanya satu orang atau 1,3 persen yang bekerja sebagai nelayan baik orang tua laki-laki maupun orang tua perempuan. Hal ini menjadi perhatian penulis tentang bagaimana pendidikan untuk anak-anak nelayan. Hasil analisis deskriptif frekuensi untuk variabel pekerjaan orang tua mahasiswa Mentawai dapat dilihat pada tabel berikut ini:

Tabel 4. Pekerjaan Orang Tua Mahasiswa

NO	Pekerjaan	Ibu	presentasi	Ayah	Presentasi
1	Ibu rumah tangga	18	23,7	0	0
2	Petani	37	48,7	25	68,4
3	Nelayan	1	1,3	1	1,3
4	Pedagang	8	10,5	5	6,6
5	Wiraswasta	5	6,6	9	11,8
6	Pegawai Negeri	5	6,6	7	9,2
7	Pegawai swasta	1	1,3	1	1,3
8	Lainya	1	1,3	1	1,3
	Total	76	100	76	100

3. Kebutuhan Kuliah Biaya konsumsi per bulan Tabel berikut ini mendeskripsikan tentang kebutuhan kuliah mahasiswa di Provinsi Sulawesi Tenggara.

Tabel 5. Biaya Konsumsi Mahasiswa

NO	Biaya Konsumsi	Frekuensi	Presentasi
1	Kurang dari Rp 299.000	10	13,2
2	Rp 300.000 – Rp 399.000	23	30,3
3	Rp 400.000 – Rp 499.000	15	19,7
4	Rp 500.000 – Rp 599.000	19	25
5	Rp 600.000 – Rp 699.000	2	2,6
6	Rp 700.000 – Rp 799.000	2	2,6
7	Rp 700.000 – Rp 799.000	5	6,6
	Total	100	100

3.5 Status Pergaulan Teman Mahasiswa

Proses pengambilan data pergaulan teman mahasiswa diambil dari Quisioner yang dibagikan dengan jumlah pertanyaan terkait pergaulan sebanyak 25 pertanyaan dan data yang akan diclusterkan adalah jumlah dari hasil jawaban quisioner, dengan nilai jawaban sebagai berikut:

- STS : Sangat Tidak Sesuai (1)
- TS : Tidak Sesuai (2)
- N : Netral (3)
- S : Sesuai (4)
- SS : Sangat Sesuai (5)

NO	NAMA MAHASISWA	p1	p2	p3	p4	p5	p6	p7	p8	p9	p10	p11	p12	p13	p14	p15	p16	p17	p18	p19	p20	p21	p22	p23	p24	p25	JUMLAH
1	WA ODE NURHAJJA	5	4	5	5	5	4	4	4	5	4	4	4	5	4	4	4	4	3	5	3	4	5	5	4	108	
2	NASRI	5	4	5	4	5	5	4	5	3	4	5	3	4	5	3	5	5	5	5	5	5	4	4	4	111	
3	EVI PURNAMA SARI	5	4	4	5	3	5	5	5	4	5	4	4	5	5	4	4	5	5	5	5	5	4	4	4	112	
4	SARTIKA	5	4	4	5	5	5	5	5	4	4	4	4	5	5	5	3	5	4	5	4	4	4	4	4	110	
5	HEPIN	5	4	4	4	4	4	5	5	5	5	4	4	2	3	5	5	5	5	2	4	4	4	5	5	108	
6	YUSRI YANSA	5	3	4	4	5	4	5	5	5	5	4	4	4	4	5	5	5	5	5	4	4	5	5	4	112	
7	ERLITA	5	5	4	4	5	4	5	5	5	5	4	4	4	5	5	5	5	5	5	4	4	5	5	4	115	
8	MUHAMMAD FEBRIAN MUNIR	5	5	4	4	5	4	5	5	5	5	4	3	3	4	4	4	4	4	4	4	4	5	5	4	103	
9	SALHUDIN HAULIN SANGGA	5	5	4	4	5	4	4	4	5	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	105	
10	LISMA HIDAYAH HALIM	5	5	4	4	4	5	4	4	5	5	4	4	4	5	5	5	5	5	2	5	5	5	5	5	114	
11	NOVITA WULANDARI	5	5	4	4	4	5	4	4	5	5	4	4	4	5	4	4	4	4	4	4	4	4	4	4	106	
12	NOVAL NIATI	5	5	4	4	4	5	4	4	5	5	4	4	4	5	5	5	5	5	5	5	5	5	5	5	117	
13	VIVIT ITMIFATIHA	5	5	4	4	4	5	4	4	5	5	4	4	5	4	4	4	4	4	4	4	4	5	5	5	110	
14	LA ODE NARWI	5	5	4	3	4	4	4	4	4	5	5	2	4	1	4	1	2	4	2	4	3	4	5	4	92	
15	MUH. DENNY P	5	5	5	5	4	3	5	4	4	5	5	5	5	4	4	5	5	3	2	5	5	4	5	4	111	
16	ARDI IRWANTO	5	5	5	5	4	4	5	4	4	5	5	5	4	4	4	5	5	4	5	5	4	5	4	5	115	
17	DEVIKA MUTMAINA	5	5	5	5	4	4	5	4	4	5	5	5	4	4	2	5	4	4	5	4	5	4	5	4	112	
18	NOKLELI	5	5	5	5	4	4	5	4	4	5	5	5	4	4	3	5	4	4	5	5	4	5	4	5	113	
19	ARSAD	5	4	5	5	4	4	5	4	4	5	5	5	4	4	4	5	4	4	4	5	4	5	4	5	113	
20	AJI PRATOMO	5	4	5	5	4	4	5	4	4	5	5	5	4	4	4	5	4	2	5	5	4	5	4	5	111	
21	BUSRAN MATAPA	5	4	5	5	4	4	5	4	4	5	5	5	4	4	4	5	4	5	5	5	4	5	4	4	113	
22	MUH. ARDIJH SYAFIT	5	4	5	5	4	4	5	4	4	5	4	4	5	5	4	4	4	4	5	5	5	4	5	4	112	

(Sumber : Hasil Perhitungan jawaban Quisioner mahasiswa STMIK Bina Bangsa Kendari)

Gambar 6. Hasil quisioner

3.6 Proses Data Mining

Setelah melakukan proses transformasi data, maka akan dilanjutkan ke proses selanjutnya yaitu proses mining data. Data mining merupakan proses mencari pola atau informasi menarik dalam data yang terpilih dengan menggunakan teknik atau metode tertentu. Pemilihan teknik dan algoritma yang tepat sangat bergantung pada proses KDD secara keseluruhan. Pada penelitian ini penerapan data mining menggunakan teknik clustering dan algoritma K-Means, Sebagai berikut;

NO	NAMA MAHASISWA	STAMBUK	NILAI PERGAULAN TEMAN	NILAI PENDAPATAN ORANG TUA/MAHASISWA	NILAI PRESTASI BELAJAR SISTEM OPERASI
1	WA ODE NURHAJJA	201751083	108	3000000	90
2	NASRI	201751017	111	2000000	106
3	EVI PURNAMA SARI	201751085	112	1000000	78
4	SARTIKA	201751044	110	500000	94
5	HEPIN	201751075	108	1000000	90
6	YUSRI YANSA	201751008	112	1000000	75
7	ERLITA	201751076	115	1000000	78
8	MUHAMMAD FEBRIAN MUNIR	201751002	103	3000000	97
9	SALHUDIN HAULIN SANGGA	201751021	105	3000000	92
10	LISMA HIDAYAH HALIM	201751014	114	4000000	81
11	NOVITA WULANDARI	201751025	106	1000000	75
12	NOVAL NIATI	201751019	117	3000000	70
13	VIVIT ITMIFATIHA	201751013	110	2500000	71
14	LA ODE NARWI	201751007	92	500000	19
15	MUH. DENNY P	201751001	111	2500000	19
16	ARDI IRWANTO	201750995	115	1000000	12
17	DEVIKA MUTMAINA	201751085	112	700000	26
18	NOKLELI	201751022	113	700000	21
19	ARSAD	201751058	113	1000000	76
20	AJI PRATOMO	201751094	111	500000	73

Gambar 7. Hasil clustering data

Proses Clustering menggunakan aplikasi Rapid Miner menjadi 3 Cluster dimana cluster 2 mendapatkan hasil sangat baik dibuktikan dengan pergaulan yang sangat baik dan status ekonomi sangat baik, cluster 0 mendapatkan hasil baik sedangkan cluster 1 menghasilkan nilai kurang baik, dengan representasi sebagai berikut;

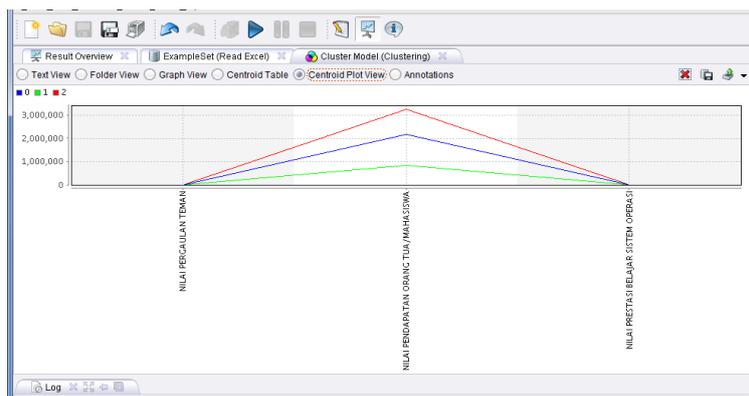
1. Pergaulan ≥ 112 sangat baik, ≥ 100 baik dan < 100 kurang baik.
2. Status ekonomi $\geq 3.000.000$ menengah atas, $\geq 2.000.000$ menengah, $< 2.000.000$ menengah kebawah.
3. Nilai Mahasiswa ≥ 95 sangat memuaskan, ≥ 80 memuaskan, < 80 kurang memuaskan.

Dengan hasil sebagai berikut;

Row No.	KODE MAH...	cluster	NILAI PERG...	NILAI PEND...	NILAI PRES...
1	M1	cluster_2	108	3000000	90
2	M2	cluster_0	111	2000000	106
3	M3	cluster_1	112	1000000	78
4	M4	cluster_1	110	500000	94
5	M5	cluster_1	108	1000000	90
6	M6	cluster_1	112	1000000	75
7	M7	cluster_1	115	1000000	78
8	M8	cluster_2	103	3000000	97
9	M9	cluster_2	105	3000000	92
10	M10	cluster_2	114	4000000	81
11	M11	cluster_1	106	1000000	75
12	M12	cluster_2	117	3000000	70
13	M13	cluster_0	110	2500000	71
14	M14	cluster_1	92	500000	19
15	M15	cluster_0	111	2500000	19
16	M16	cluster_1	115	1000000	12
17	M17	cluster_1	112	700000	26

Gambar 8. Hasil clustering rapid miner

Hasil grafik Centroid Plat View sebagai berikut, cluster 2 warna merah menunjukkan grafik tertinggi dengan penghasilan diatas 3 juta, cluster 0 warna biru mendapatkan nilai sedang dan cluster 1 dengan hasil paling rendah:



Gambar 9. Hasil Centroid Plat View

Jumlah data yang dicluster sebanyak 76 dengan cluster model, cluster 2 sebanyak 17 item, cluster 0 sebanyak 17 item sedangkan cluster 1 sebanyak 42 item.

```

Cluster Model

Cluster 0: 17 items
Cluster 1: 42 items
Cluster 2: 17 items
Total number of items: 76
    
```

Gambar 10. Text View Cluster Model

4. Kesimpulan dan saran

4.1 Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah dijelaskan dapat disimpulkan bahwa perancangan dan Implementasi Aplikasi data mining untuk mengelompokan prestasi belajar mahasiswa didasarkan pada status ekonomi keluarga sangat berpengaruh khususnya jika kuliah dibidang komputer. Mahasiswa secara tidak langsung membutuhkan Komputer dan peralatan lainya untuk menunjang kebutuhan kuliah agar tidak ada kendala dalam proses belajarnya, oleh karena itu status ekonomi sangat berpengaruh dan juga pergaulan dengan kelompok mahasiswa yang lebih mementingkan belajar dibandingkan kegiatan lainnya yang tidak bermanfaat bagi prestasi belajar mahasiswa. Oleh sebab itu bergaul dengan teman mahasiswa yang rajin belajar juga status ekonomi dapat mempengaruhi prestasi belajar mahasiswa.

4.2 Saran

Berdasarkan data hasil penelitian dan pembahasan yang telah dikemukakan maka peneliti dapat memberikan saran-saran sebagai berikut:

1. Bagi mahasiswa
Lingkungan pergaulan teman memiliki peranan dalam meningkatkan hasil belajar, mahasiswa diharapkan lebih selektif dalam memilih pergaulan teman sehingga dapat mendukung kegiatan pembelajaran dan mahasiswa tidak terpengaruh hal-hal negatif yang dapat merugikan diri sendiri terutama terkait dengan hasil belajar mahasiswa.
2. Bagi peneliti selanjutnya
Untuk penelitian selanjutnya dapat dilakukan dengan data yang lebih banyak dan parameter yang lebih banyak lagi, sehingga memungkinkan mendapatkan metode dan fokus penelitian jangan hanya pada pergaulan teman dan status ekonomi yang tentunya dapat meningkatkan prestasi belajar mahasiswa.
3. Gunakan metode yang lain atau lebih dari satu metode dengan harapan mendapatkan parameter hasil yang lebih baik dalam peningkatan belajar mahasiswa sehingga nantinya dapat dijadikan referensi dosen dalam meningkatkan kualitas mahasiswa.

Daftar Pustaka

- [1] S. M. Dr. Suyanto, *Data Mining Untuk Klasifikasi Dan Klasterisasi Data*. Bandung: Informatika, 2017.
- [2] R. T. Vlandari, "Pengelompokan Tingkat Keamanan Wilayah Jawa Tengah Berdasarkan Indeks Kejahatan Dan Jumlah Pos Keamanan Dengan Metode Klastering K-Means," *J. Ilm. SINUS*, vol. Vol 14, No, no. ISSN :1693-1173, pp. 59–72, 2016.
- [3] J. O. Ong, "Implementasi Algoritma K-means clustering untuk menentukan strategi marketing president university," *J. Ilm. Tek. Ind.*, vol. vol.12, no, no. juni, pp. 10–20, 2013.
- [4] E. Prasetyo, *Data Mining Konsep dan Aplikasi menggunakan MatLab*. Yogyakarta: Andi, 2012.
- [5] Ilma mulwafa, *Psikologi Remaja Terhadap Pergaulan Serta Perkembangan Peserta Didik*. Jakarta: PT Bumi Aksara.
- [6] T. H. Handoko, *Manajemen Personalia dan Sumber Daya Manusia*. Yogyakarta: BPF, 2012.
- [7] S. dan S. Soekanto, *Sosiologi Suatu Pengantar*. Jakarta: Raja Grafindo Persada, 2013.