

**IMPLEMENTASI ALGORITMA SUPPORT VECTOR MACHINE (SVM) DALAM
MEMPREDIKSI KELULUSAN JALUR MASUK PERGURUAN TINGGI
BANDA ACEH
(STUDI KASUS MAHASISWA BARU TAHUN AJARAN 2019)**

**SKRIPSI
Diajukan Oleh:**

**MARHAMAH
NIM. 160212034**

**Mahasiswa Fakultas Tarbiyah dan Keguruan (FTK)
Prodi Pendidikan Teknologi Informasi**



**FAKULTAS TARBIYAH DAN KEGURUAN
UNIVERSITAS ISLAM NEGERI AR-RANIRY
DARUSSALAM-BANDA ACEH
2020 M/ 1441 H**

**IMPLEMENTASI ALGORITMA SUPPORT VECTOR MACHINE (SVM) DALAM
MEMPREDIKSI KELULUSAN JALUR MASUK PERGURUAN TINGGI BANDA
ACEH (STUDY KASUS; MAHASISWA BARU TAHUN AJARAN 2019)**

SKRIPSI

Diajukan Kepada Fakultas Tarbiyah dan Keguruan (FTK)
Universitas Islam Negeri Ar-Raniry Darussalam Banda Aceh
Sebagai Beban Studi Untuk Memperoleh Gelar Sarjana
Dalam Ilmu Pendidikan Teknologi Informasi

Oleh

MARHAMAH
NIM. 16021203

Mahasiswa Fakultas Tarbiyah dan Keguruan
Prodi Pendidikan Teknologi Informasi

Disetujui Oleh:

Pembimbing I



Bustami, MSc

NIP. 19860408 201403 1 001

Pembimbing II



Hendri Ahmadian, S.Si., M.I.M

NIP. 19830104 201403 1 002

**IMPLEMENTASI ALGORITMA SUPPORT VECTOR MACHINE (SVM) DALAM
MEMPREDIKSI KELULUSAN JALUR MASUK PERGURUAN TINGGI
BANDA ACEH
(STUDI KASUS MAHASISWA BARU TAHUN AJARAN 2019)**

SKRIPSI

**Telah Diuji Oleh Panitia Ujian Munaqasyah Skripsi
Fakultas Terbiyah dan Keguruan UIN Ar-Raniry dan Dinyatakan Lulus
Serta Diterima Sebagai Salah Satu Beban Studi Program Sarjana (S-1)
Dalam Ilmu Pendidikan Teknologi Informasi**

Pada Hari/Tanggal

Selasa,

18 Agustus 2020

28 Dzulhijah 1441 H

Panitia Ujian Munaqasyah Skripsi

Ketua

Bustami, M.Sc

NIP. 19860408 201403 1 001

Sekretaris

Izzah Al-Fikry, M.Pd

Penguji I,

Hendri Ahmadian, S.Si., M.I.M

NIP.19830104 201403 1 002

Penguji II,

Zuhra Sofyan, M.Sc

NIP.19840309 20181 1 001

Mengetahui,
Dekan Fakultas Terbiyah dan Keguruan
Darussalam - Banda Aceh



Dr. Mubliq Rizali, S.H., M.Ag

NIP. 195906091989031001

LEMBAR PERNYATAAN KEASLIAN KARYA ILMIAH

Yang bertanda tangan di bawah ini, saya:

Nama : Marhamah

NIM : 160212034

Program Studi : Pendidikan Teknologi Informasi

Fakultas : Tarbiyah dan Keguruan

Judul Skripsi : Implementasi Algoritma Support Vector Machine (SVM)

Dalam Memprediksi Kelulusan Jalur Masuk Perguruan Tinggi
Banda Aceh (Study Kasus; Mahasiswa Baru Tahun Ajaran
2019).

Dengan ini menyatakan bahwa dalam penulisan skripsi ini, saya:

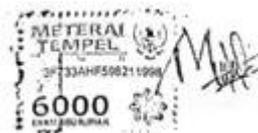
1. Tidak menggunakan ide orang lain tanpa mampu mengembangkan dan mempertanggung jawabkannya.
2. Tidak melakukan plagiasi terhadap naskah karya orang lain.
3. Tidak menggunakan karya orang lain tanpa menyebutkan sumber asli atau tanpa izin pemiliknya.
4. Tidak memanipulasi dan memalsukan data.
5. Mengerjakan sendiri karya ini dan mampu bertanggung jawab atas karya ini.

Bila dikemudian hari ada tuntutan dan pihak lain atas karya saya, dan telah melalui pembuktian yang dapat bertanggung jawabkan dan ternyata memang ditemukan bukti bahwa saya telah melanggar persyaratan, amaka saya siap dikenai sanksi berdasarkan aturan yang berlaku di Fakultas Tarbiyah dan Keguruan UIN Ar-Raniry Banda Aceh.

Dengan demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari pihak manapun.

Banda Aceh, 18 Agustus 2020

Yang menyatakan,



Marhamah

NIM. 160212034

ABSTRAK

Nama : Marhamah
NIM : 160212034
Fakultas/Prodi : Tarbiyah dan Keguruan/Pendidikan Teknologi Informasi
Judul : Implementasi Algoritma Support Vector Machine (SVM) Dalam Memprediksi Kelulusan Jalur Masuk Perguruan Tinggi Banda Aceh (Studi Kasus Mahasiswa Baru Tahun Ajaran 2019).
Tanggal Sidang : 18 Agustus 2020
Tebal Skripsi : 65 Halaman
Pembimbing I : Bustami, M.Sc
Pembimbing II : Hendri Ahmadian,S.Si.,M.I.M

Pendidikan tinggi merupakan salah satu lembaga untuk mempersiapkan peserta didik menjadi anggota masyarakat yang bisa memenuhi kebutuhan masyarakat dalam berbagai bidang. Hal tersebut menyangkut dengan perubahan UU pendidikan tinggi tentang penerimaan mahasiswa baru. Saat ini ada beberapa jenis jalur masuk penerimaan mahasiswa baru di perguruan tinggi. Banyak cara yang bisa dipersiapkan oleh setiap calon mahasiswa untuk bisa lulus ujian seleksi masuk PTN. Oleh karena itu peneliti ingin memprediksi kelulusan jalur masuk perguruan tinggi dengan metode klasifikasi menggunakan algoritma SVM yang dibantu dengan *machine learning* WEKA, menggunakan data mahasiswa baru tahun ajaran 2019. Hasil akhir pada penelitian ini terdapat dua variabel yang memiliki hubungan paling baik yaitu variabel bintel seleksi masuk PTN dan variabel jalur minat, dengan variabel bintel seleksi masuk PTN memiliki nilai *pearson correlation* sebesar -0,180** dan nilai signifikansi sebesar 0,002, jalur minat memiliki nilai akurasi sebesar 0,311** dan nilai signifikansi sebesar 0,000. Kemudian berdasarkan hasil pengujian *cross-validation* dan *percentage split* algoritma SVM memiliki akurasi yang sangat baik dengan rata-rata akurasi mencapai 99% dengan nilai AUC (*Area Under Curve*) sebesar 0.9907 dan waktu konsumsi yang dibutuhkan untuk pengujian hanya sebesar 0.01-0.02 *second*.

Keyword : Jalur Masuk PTN, *Data Mining*, Klasifikasi ,Evaluasi, WEKA, SVM, *Cross-Validation*, *Percentage Split*.

KATA PENGANTAR



Alhamdulillahirabil'alamin puji dan syukur kita panjatkan atas kehadiran Allah SWT atas segala rahmat dan hidayah-Nya, Shalawat beriring salam untuk tuntunan suri tauladan Rasulullah Shalallahu'alaihiwasalam beserta seluruh keluarga dan sahabat beliau yang senantiasa menjunjung tinggi nilai-nilai Islam serta menggali ilmu yang tiada habisnya yang sampai saat ini masih dapat dinikmati oleh seluruh manusia dipenjuru dunia, sehingga penulis dapat menyelesaikan skripsi ini dengan judul "*Implementasi Algoritma SVM Dalam Memprediksi Kelulusan Jalur Masuk Perguruan Tinggi Banda Aceh*".

Penulisan skripsi ini merupakan salah satu syarat untuk mendapatkan gelar sarjana pada Fakultas Tarbiyah dan Keguruan UIN Ar-Raniry Darussalam Banda Aceh. Dalam usaha penyusunan skripsi ini, penulis banyak sekali menghadapi kesulitan teknik penulisan maupun dalam penguasaan bahan. Walaupun demikian, penulis tidak putus asa dalam berusaha dan dengan adanya dukungan dari berbagai pihak, terutama sekali dosen pembimbing, kesulitan tersebut dapat teratasi. Pada kesempatan ini, penulis mengucapkan ribuan terima kasih kepada :

1. Kedua orang tua ayahanda tercinta dan ibunda tersayang yang telah memberikan dukungan moril maupun materil serta doa yang tiada henti kepada penulis.
2. Segenap keluarga dan sahabat saya Rike Mahara yang selalu menyemangati dan membantu dari awal hingga akhir penyelesaian skripsi ini.
3. Bapak Rektor UIN Ar-Raniry, Prof. Dr. H. Warul Walidin AK. MA yang selalu mendukung kami dan memberi motivasi.
4. Bapak Dekan Fakultas Tarbiyah dan Keguruan UIN Ar-Raniry, bapak Dr. Muslim Razali, S.H., M.Ag.
5. Bapak Bustami, M.Sc selaku pembimbing pertama dan Bapak Hendri Ahmadian, S.Si., M.I.M selaku pembimbing kedua yang telah meluangkan waktunya dan mencurahkan pemikirannya dalam membimbing penulis untuk menyelesaikan karya ilmiah ini.
6. Ketua Prodi Pendidikan Teknologi Informasi (PTI) bapak Yusran, M.Pd., Sekretaris Prodi Pendidikan Teknologi Informasi bapak Hazrullah, M.Pd, serta staf Prodi yang telah banyak membantu proses pelaksanaan penelitian untuk penulisan skripsi ini.
7. Bapak/ibu dosen pengajar Program Studi Pendidikan Teknologi Informasi yang telah membekali penulis dengan berbagai ilmu pengetahuan sehingga dapat menyelesaikan studi ini.
8. Pihak Urusan Alumni Fakultas Tarbiyah dan Keguruan, serta para alumni yang telah membantu proses pelaksanaan penelitian untuk skripsi ini.

9. Sahabat dan teman-teman mahasiswa Jurusan Pendidikan Teknologi Informasi leting 2016 serta seluruh keluarga PTI yang telah mendoakan dan memberi dukungan selama ini.
10. Dan untuk semuanya yang tidak dapat penulis sebutkan satu persatu. Penulis berserah diri kepada Allah karena tidak ada yang terjadi tanpa kehendak-Nya. Segala usaha telah dilakukan untuk menyempurnakan skripsi ini. Namun, penulis menyadari dalam penulisan skripsi ini masih banyak ditemukan kekurangan dan kekhilafan. Oleh karena itu, penulis mengharapkan saran yang dapat dijadikan masukan guna perbaikan di masa yang akan datang. Semoga Allah SWT meridhai penulisan ini dan senantiasa memberikan rahmat dan hidayah-Nya kepada kita semua. Amin ya rabbal' alamin.



Banda Aceh, 4 Juli 2020

Penulis,

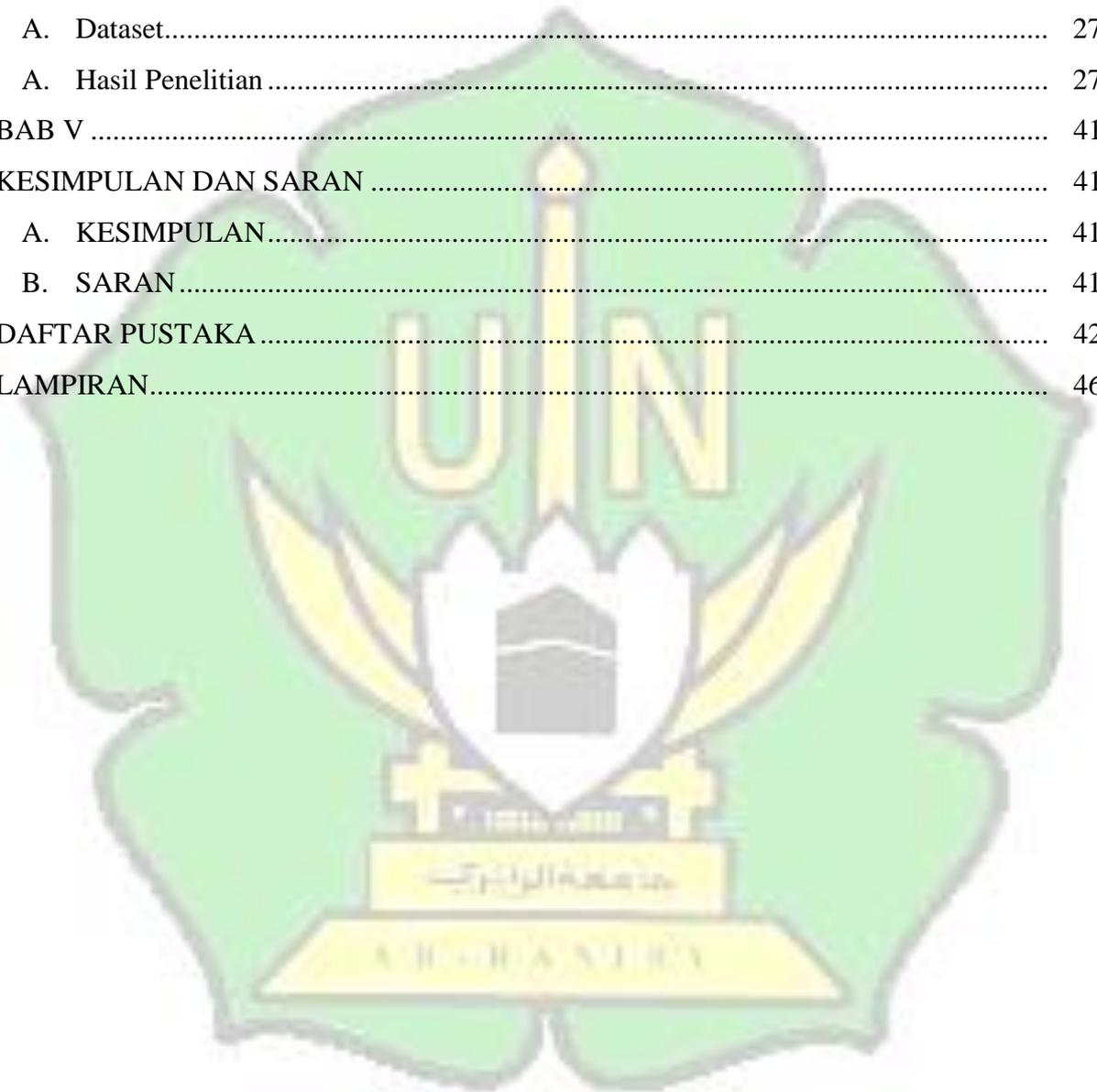
Marhamah

NIM.160212034

DAFTAR ISI

	Halaman
LEMBARAN JUDUL	
PENGESAHAN PEMBIMBING	
PENGESAHAN SIDANG	
LEMBAR PERNYATAAN KEASLIAN	
ABSTRAK.....	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vii
DAFTAR GAMBAR.....	ix
DAFTAR TABEL.....	x
DAFTAR GRAFIK.....	xi
BAB 1	1
PENDAHULUAN	1
A. Latar Belakang	1
B. Rumusan Penelitian	3
C. Tujuan Penelitian	3
D. Batasan Penelitian.....	3
E. Manfaat Penelitian	3
BAB II.....	4
TINJAUAN PUSTAKA	4
A. Jalur Masuk Perguruan Tinggi.....	4
B. Data Mining	5
C. Klasifikasi	7
D. SVM.....	8
E. <i>Naive Bayes</i>	11
F. <i>Bayes Network</i>	11
G. <i>Decision Tree</i>	11
H. <i>Weka</i>	12
I. Metode Evaluasi.....	12
J. Penelitian Terdahulu	14
BAB III	16
METODOLOGI PENELITIAN	16

A. Spesifikasi Perangkat	16
B. Tempat dan Waktu Penelitian	16
C. Teknik Pengumpulan Data.....	16
D. Prosedur Penelitian	16
BAB IV	27
HASIL DAN PEMBAHASAN	27
A. Dataset.....	27
A. Hasil Penelitian	27
BAB V	41
KESIMPULAN DAN SARAN	41
A. KESIMPULAN.....	41
B. SARAN.....	41
DAFTAR PUSTAKA.....	42
LAMPIRAN.....	46



DAFTAR GAMBAR

No. Gambar	Halaman
Gambar 2. 1 Tahap-tahap data mining[21]	6
Gambar 2. 2 SVM mencoba menemukan hyperplane terbaik untuk	8
Gambar 2. 3 Temukan fungsi pemisahan optimal untuk objek	9
Gambar 3. 1 Prosedur penelitian.....	17
Gambar 3. 2 Sampel data dalam format CSV.....	20
Gambar 3. 3 Sampel data dalam format Arff.....	20
Gambar 3. 4 Flowchart SVM.....	23
Gambar 3. 5 Visualisasi garis hyperplane.....	25
Gambar 4. 1 Korelasi jalur minat dengan jalur lulus	28
Gambar 4. 2 Korelasi bimbel seleksi masuk PTN dengan jalur lulus	29
Gambar 4. 3 Korelasi nilai UN dengan jalur lulus.....	29
Gambar 4. 4 Korelasi organisasi di SMA dengan jalur lulus	30
Gambar 4. 5 Korelasi les private di SMA dengan jalur lulus	30
Gambar 4. 6 Korelasi pekerjaan ayah dengan jalur lulus	31
Gambar 4. 7 Korelasi pekerjaan ibu dengan jalur lulus.....	31
Gambar 4. 8 Korelasi pendidikan terakhir ayah dengan jalur lulus.....	32
Gambar 4. 9 Korelasi pendidikan terakhir ibu dengan jalur lulus	32
Gambar 4. 10 Korelasi orang tua yang dimiliki dengan jalur lulus	33
Gambar 4. 11 Perbandingan Jenis Pengujian.....	38
Gambar 4. 12 Kurva ROC dengan Algoritma SVM.....	38

DAFTAR TABEL

No. Tabel	Halaman
Tabel 2. 1 Model Confusion matrix untuk masalah klasifikasi kelas yang diprediksi[20]...	12
Tabel 3. 1 Variabel kuesioner	18
Tabel 3. 2 Data asli	19
Tabel 3. 3 Data sesudah cleaning dan korelasi	19
Tabel 3. 4 Contoh confussion matrik	26
Tabel 4. 1 Pemberian nama kelas pada label	27
Tabel 4. 2 Hubungan korelasi	27
Tabel 4. 3 Kategori Kekuatan Korelasi.....	28
Tabel 4. 4 Perbandingan Korelasi Variabel	33
Tabel 4. 5 Perbanding uji variabel	34
Tabel 4. 6 Cross Validation	35
Tabel 4. 7 Percentage Split	37
Tabel 4. 8 Accuracy dan AUC.....	39
Tabel 4. 9 Perbandingan akurasi.....	40
Tabel 4. 10 Waktu konsumsi	40
Tabel 4. 11 Perbanding uji variabel	54

DAFTAR GRAFIK

Grafik 4. 1 Perbandingan akurasi pada variabel asli, variabel sesudah cleaning, variabel korelasi dan tanpa variabel korelasi.....	35
Grafik 4. 2 Cross Validation	36
Grafik 4. 3 Percentage Split	37



DAFTAR LAMPIRAN

No.Lampiran

1. Surat keputusan dekan
2. Data
3. Praposes pada weka
4. Pengujian cross-validation dan percentage split
5. Akurasi dan confussion matrix
6. Bentuk data dalam binary
7. Kurva ROC
8. Uji korelasi
9. Pengujian akurasi
10. Uji evaluasi
11. Waktu konsumsi



BAB 1

PENDAHULUAN

A. Latar Belakang

Pendidikan tinggi adalah salah satu lembaga menyiapkan peserta didik untuk menjadi anggota masyarakat yang bisa memenuhi kebutuhan masyarakat dalam berbagai bidang[1]. Hal tersebut menyangkut dengan perubahan UU pendidikan tinggi tentang penerimaan mahasiswa baru[2]. Mengenai proses penerimaan mahasiswa di Perguruan Tinggi terdapat beberapa jalur masuk yaitu SNMPTN,SBMPTN,UM-PTKIN,SPAN-PTKIN dan PMB sebagaimana yang telah diterangkan dalam peraturan pemerintah RI Nomor 34 tahun 2010[3].

Saat ini ada beberapa jenis jalur masuk penerimaan mahasiswa baru di perguruan tinggi diantaranya yaitu Seleksi Nasional Masuk Perguruan Tinggi Negeri (SNMPTN) yang dipilih berdasarkan hasil pencarian akademik yang dicapai oleh calon mahasiswa[4]. Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN) yaitu seleksi berdasarkan hasil tes tertulis menggunakan metode pencetakan (test berbasis kertas) atau kombinasi dari hasil tes tertulis dan tes kemampuan yang dimiliki[5]. Selanjutnya SPAN-PTKIN merupakan Seleksi yang dilakukan secara nasional oleh semua universitas agama Islam seperti UIN / IAIN / STAIN dalam sistem yang terintegrasi dan juga dilakukan oleh komite eksekutif yang dibentuk oleh Menteri Agama Republik Indonesia[6].Sedangkan Jalur Mandiri merupakan seleksi yang dilakukan secara independen oleh masing-masing universitas negeri dengan memanfaatkan skor hasil tes SBMPTN[7]. UMPTKIN merupakan jalur masuk perguruan tinggi yang dilakukan secara nasional oleh semua Perguruan Tinggi Islam Negeri seperti UIN, IAIN dan STAIN dalam sistem yang terintegrasi yang dilakukan secara bersama atas dasar peraturan Departemen Agama Republik Indonesia[8]. Terakhir ada PMB merupakan jalur mandiri penerimaan mahasiswa baru atau disebut sebagai PMB lokal. Jalur PMB ini merupakan jalur terakhir setelah menerima mahasiswa melalui jalur

SNMPTN, SPAN PTKIN, SBMPTN dan UM PTKIN [9]. Pelaksanaan penerimaan mahasiswa baru melalui jalur PMB dilakukan melalui ujian tertulis (*paper based testing*).

Banyak cara yang bisa dipersiapkan oleh setiap calon mahasiswa untuk bisa lulus ujian seleksi masuk PTN diantaranya adalah dengan mengikuti bimbel, les privat, mempunyai prestasi akademik, dan nilai UN yang memadai. Akan tetapi ada kasus dimana calon mahasiswa sudah melakukan persiapan tersebut namun hasilnya mereka dinyatakan tidak lulus. Ada juga kasus dimana calon mahasiswa tersebut tidak melakukan persiapan yang matang dinyatakan lulus di jalur yang diminati. Ini menandakan persiapan-persiapan tersebut belum bisa menjamin calon mahasiswa untuk lulus di jalur yang diminati. Dengan demikian penulis ingin meneliti penyebab utama yang mempengaruhi calon mahasiswa bisa diterima di jalur yang diminati. Metode data mining adalah salah satu cara untuk menganalisis masalah tersebut.

Berbagai metode data mining yang dapat digunakan untuk memprediksi kelulusan jalur masuk perguruan tinggi bagi calon mahasiswa di antaranya yaitu *Naive Bayes*, *Bayes Network*, *Decision Tree*, *k-nearest Neighbor* dan *SVM*. Dalam proses pengolahan data dengan menggunakan data mining, telah banyak dilakukan penelitian sebelumnya, diantaranya penelitian tentang “implementasi data mining untuk memprediksi kelulusan mahasiswa menggunakan metode *Naive Bayes*, dengan hasil akhir memiliki nilai akurasi sebesar 94% [10]. Selain itu penelitian tentang “Analisis kinerja metode *naive bayes* dan *SVM* untuk penentuan pola kelompok penyakit”, hasil yang didapatkan dengan metode *SVM* dengan nilai akurasi mencapai 99%, dan metode *Naive Bayes* dengan nilai akurasi mencapai 93%. Dari nilai akurasi yang didapatkan menunjukkan metode *SVM* lebih akurat daripada metode *naive bayes* [11]. Selanjutnya penelitian yang berhubungan dengan prediksi yang disajikan oleh Andri dengan menggunakan Algoritma yang berbeda yaitu algoritma *J48 (Decision Tree)*”. Hasil dari penelitian tersebut didapatkan bahwa ada beberapa atribut yang berpengaruh dalam menentukan

kelulusan mahasiswa yaitu tempat lahir, pekerjaan orang tua, asal sekolah dan jenis kelamin[12].

Setiap metode yang digunakan dalam penelitian ini memiliki kelemahan dan kelebihan sendiri, diantaranya akurasi rendah yang dicapai dengan setiap metode yang digunakan. Hanya metode SVM yang lebih akurat dari pada metode lainnya. Untuk mencapai pengembangan penelitian yang lebih baik, peneliti ingin menggunakan metode SVM untuk memprediksi kelulusan mahasiswa di jalur masuk PTN di Banda Aceh. Data yang digunakan berkisar antara 500-1000 data dari beberapa perguruan tinggi Banda Aceh.

B. Rumusan Penelitian

Berdasarkan latar belakang diatas peneliti dapat merumuskan beberapa masalah yaitu:

1. Bagaimana mengimplementasi algoritma SVM dalam memprediksi kelulusan calon mahasiswa di setiap seleksi jalur masuk perguruan tinggi.
2. Bagaimana akurasi algoritma SVM dalam memprediksi kelulusan calon mahasiswa di setiap seleksi jalur masuk perguruan tinggi.

C. Tujuan Penelitian

Berdasarkan latar belakang di atas penelitian ini bertujuan untuk:

1. Mengimplementasi algoritma SVM dalam memprediksi kelulusan jalur masuk perguruan tinggi studi kasus mahasiswa baru tahun ajaran 2019.
2. Mengetahui hasil implementasi algoritma SVM dalam memprediksi kelulusan calon mahasiswa di setiap seleksi jalur masuk perguruan tinggi.

D. Batasan Penelitian

Untuk lebih memfokuskan penelitian, penulis membatasi masalah yang digunakan dalam penelitian ini, yaitu:

1. Data yang digunakan dalam penelitian ini adalah data mahasiswa baru tahun ajaran 2019.
2. Metode yang digunakan dalam penelitian ini adalah metode *SVM*
3. Data set yang digunakan berkisar antara 500-1000 data mahasiswa baru tahun ajaran 2019.

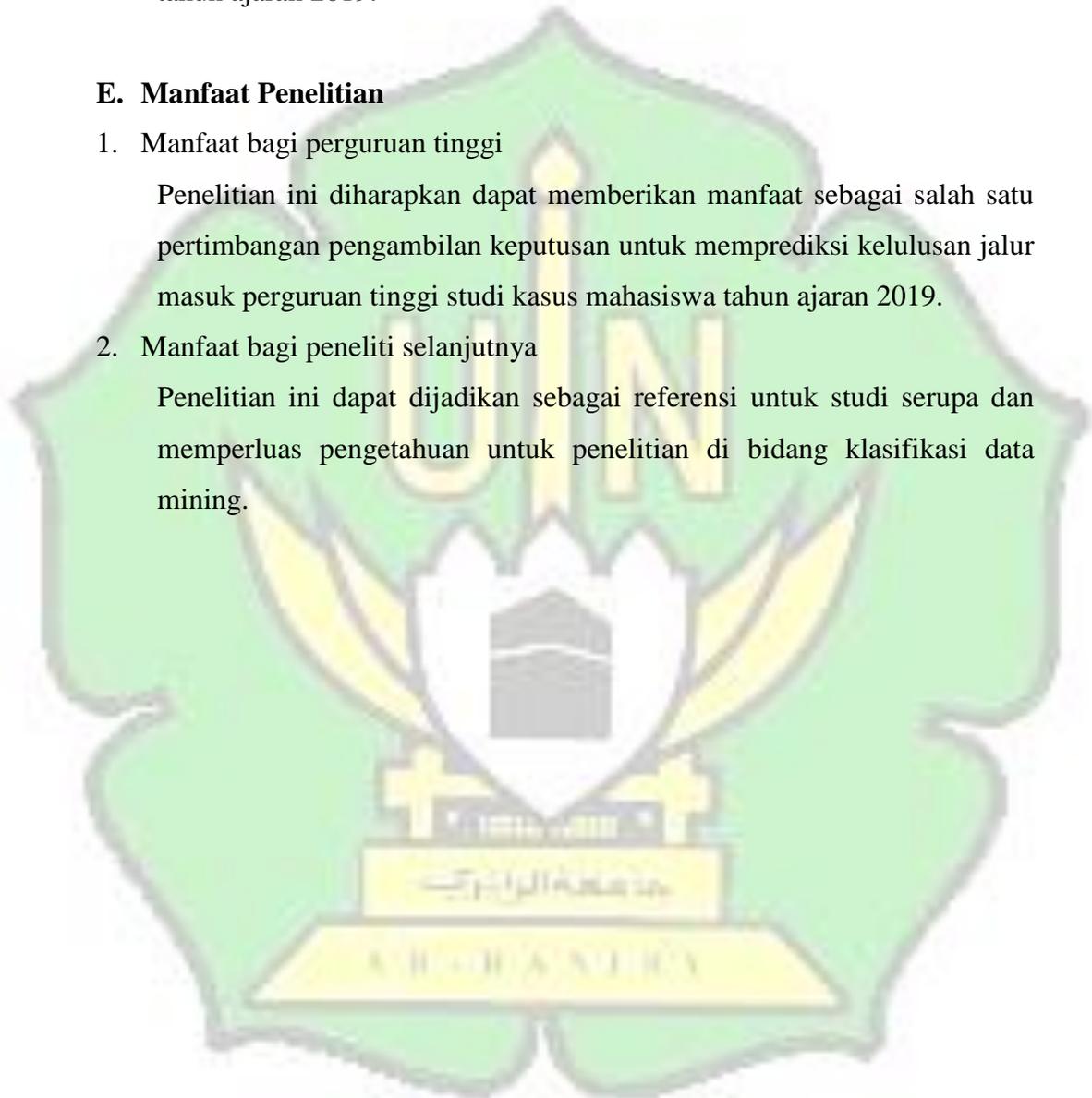
E. Manfaat Penelitian

1. Manfaat bagi perguruan tinggi

Penelitian ini diharapkan dapat memberikan manfaat sebagai salah satu pertimbangan pengambilan keputusan untuk memprediksi kelulusan jalur masuk perguruan tinggi studi kasus mahasiswa tahun ajaran 2019.

2. Manfaat bagi peneliti selanjutnya

Penelitian ini dapat dijadikan sebagai referensi untuk studi serupa dan memperluas pengetahuan untuk penelitian di bidang klasifikasi data mining.



BAB II

TINJAUAN PUSTAKA

A. Jalur Masuk Perguruan Tinggi

Pendidikan tinggi merupakan salah satu lembaga untuk mempersiapkan peserta didik menjadi anggota masyarakat yang bisa memenuhi kebutuhan masyarakat dalam berbagai bidang[1]. Hal tersebut menyangkut dengan perubahan UU pendidikan tinggi tentang penerimaan mahasiswa baru[2]. Mengenai proses penerimaan mahasiswa di Perguruan Tinggi terdapat beberapa jalur masuk yaitu SNMPTN, SBMPTN, UM-PTKIN, SPAN-PTKIN dan PMB sebagaimana yang telah diterangkan dalam peraturan pemerintah RI Nomor 34 tahun 2010[3].

1. Jalur SNMPTN

SNMPTN adalah salah satu jalur masuk perguruan tinggi yang dilaksanakan oleh masing- masing universitas, yang diseleksi berdasarkan hasil pencarian prestasi akademik yang dicapai oleh calon mahasiswa[13].

2. Jalur SBMPTN

SBMPTN merupakan jalur masuk perguruan tinggi yang seleksi berdasarkan hasil tes tertulis menggunakan metode pencetakan (tes berbasis kertas) atau kombinasi dari hasil tes tertulis dan tes kemampuan yang dimiliki yang dilaksanakan secara serentak di seluruh PTN se Indonesia [14].

3. Jalur UM-PTKIN

UMPTKIN merupakan jalur masuk perguruan tinggi yang dilakukan secara nasional oleh semua Perguruan Tinggi Islam Negeri seperti UIN, IAIN dan STAIN dalam sistem yang terintegrasi yang dilakukan secara bersama atas dasar peraturan Departemen Agama Republik Indonesia[8].

Jalur UMPTKIN adalah jalur seleksi nasional yang hampir sama dengan jalur SBMPTN karena kedua jalur tersebut dilakukan berdasarkan hasil tes tertulis di atas kertas atau menggunakan sistem komputer[15]. Perbedaannya adalah bahwa SBMPTN menerima calon mahasiswa baru untuk semua

universitas negeri di Indonesia, sementara UM-PTKIN menerima calon mahasiswa baru hanya di perguruan tinggi agama Islam di Indonesia.

4. Jalur SPAN-PTKIN

SPAN-PTKIN adalah seleksi yang dilakukan secara nasional oleh semua universitas agama Islam seperti UIN / IAIN / STAIN dalam sistem yang terintegrasi dan juga dilakukan oleh komite eksekutif yang dibentuk oleh Menteri Agama RI[6]. Jalur SPAN-PTKIN ini hampir sama dengan SNMPTN jika dilihat dari segi penyeleksiannya, perbedaannya SNMPTN berlaku untuk universitas se indonesia[8].

5. Jalur PMB

PMB merupakan jalur mandiri penerimaan mahasiswa baru atau disebut sebagai PMB lokal. Jalur PMB ini merupakan jalur terakhir setelah menerima mahasiswa beberapa jalur di atas[9]. Penerimaan mahasiswa baru melalui jalur PMB dilakukan melalui ujian tertulis (*paper based testing*).

B. Data Mining

Data mining adalah kegiatan mencari dan menggali informasi yang tidak dikenal secara manual dari database. Informasi yang dihasilkan diperoleh dengan mengekstraksi dan mengenali pola penting atau menarik dari data yang terkandung dalam database[16]. *Data mining* merupakan suatu kegiatan dengan menggunakan beberapa teknik yang bertujuan untuk mendapatkan informasi dan pengetahuan yang berhubungan dengan database besar data yang ukurannya lebih besar biasanya diolah menggunakan *data mining*, kemudian dari data tersebut dilakukan pencarian pola atau trend sesuai dengan tujuan dari penerapan *data mining* tersebut, selanjutnya hasil dari pengolahan data mining tersebut digunakan untuk pengambilan keputusan maupun hasil prediksi analisis yang dibutuhkan[17].

Data mining merupakan kegiatan analisis data dengan memanfaatkan software dan dengan menggunakan tool untuk menemukan pola dengan

mengidentifikasi aturan dan karakteristik pada database[18]. Pada dasarnya, data mining dapat dibagi menjadi dua kategori utama, yaitu[19]:

- *Descriptive mining*

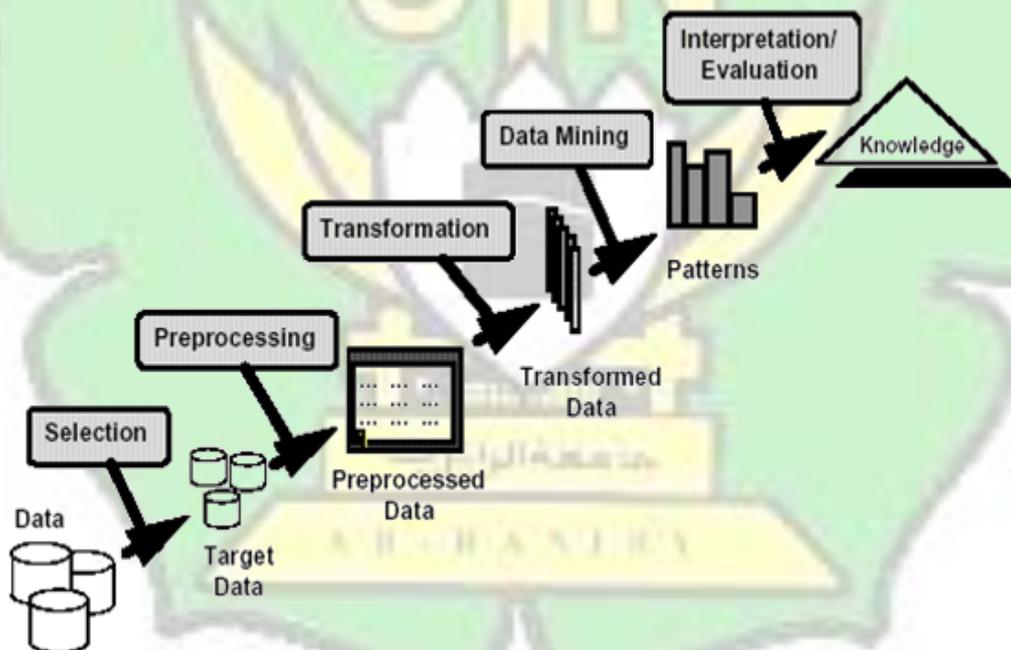
Descriptive mining merupakan kegiatan untuk menggali nilai penting dari sebuah database yang tersembunyi dan menemukan pola data tertentu yang belum diketahui sebelumnya.

- *Prediktif*

Prediktif merupakan proses pencarian pola dari data dengan menggunakan beberapa atribut lain di masa akan datang. Klasifikasi termasuk salah satu teknik yang terdapat dalam predictif mining [20].

1. Tahap-Tahap Data mining

Dalam sejumlah proses, data mining dapat dibagi menjadi beberapa fase[21] :



Gambar 2. 1 Tahap-tahap data mining[21]

1. Seleksi Data (*Data Selection*)

Dalam penelitian ini menggunakan data mahasiswa baru tahun ajaran 2019 di beberapa Universitas di Banda Aceh, yang terdiri dari data

mahasiswa dan data jalur masuk PTN. Jumlah data mahasiswa yang akan diperoleh berkisar antara 500-1000 data yang terdiri dari 22 variabel yang menjelaskan identitas mahasiswa dan informasi tentang kondisi mahasiswa yang bersangkutan. Atribut tersebut diantaranya adalah **jumlah saudara_kandung, ortu yang dimiliki, yang membiayai sekolah, pend_ayah/ibu, pek_ayah/ibu, peng_ortu, asal_sek, jur_sek, les privat_sek, prestasi_sek, organisasi _sek, nilai_UN, bimbel jalur masuk PTN, jalur yang diikuti, jalur minat, jalur_lulus, sesuai atau tidak dengan jalur minat dan jurusan_lulus**. Dari semua variabel diatas akan diseleksi untuk melihat variabel yang berpengaruh dalam penelitian ini, dengan cara mencari korelasi dari setiap atribut.

2. Pembersihan data (*Data cleaning*)

Langkah kedua dalam proses data mining adalah membersihkan data-data yang tidak mempunyai nilai yang utuh, tidak konsisten atau tidak relevan, dan data yang berlebihan. Semua atribut di atas akan dipilih untuk mendapatkan atribut yang memiliki nilai relevan tidak outlier dan tidak berlebihan. Ketiga syarat tersebut yang wajib dilakukan untuk mendapatkan data yang bersih untuk diaplikasikan pada data mining[21].

3. Transformasi data (*Data Transformation*)

Transformasi data merupakan proses perubahan data. Dalam penelitian ini peneliti menggunakan data excel, data tersebut di konversi ke dalam format CSV, kemudian dikonversi lagi ke format weka yaitu arff, supaya menjadi data yang bisa diolah dalam data mining[19].

C. Klasifikasi

Klasifikasi terdiri dari dua tahapan yang merupakan salah satu metode data mining yaitu tahap pembelajaran dan tahap klasifikasi. Tahapan pembelajaran merupakan tahap membangun model klasifikasi sedangkan tahap klasifikasi

yaitu tahap mengaplikasikan model klasifikasi untuk memprediksi label kelas dari sebuah data[22]. Klasifikasi berfungsi untuk membagikan objek ke dalam beberapa kelas yang diolah untuk mempertahankan aturan klasifikasi yang digunakan untuk memprediksi label kelas yang belum diketahui[23]. Klasifikasi menemukan model baru yang dapat digunakan untuk memprediksi kelas objek dengan nama kelas yang tidak diketahui[24].

Dari pengertian diatas klasifikasi dapat disimpulkan sebagai salah satu metode dalam data mining yang berfungsi untuk menemukan model baru dari suatu data untuk memprediksi label kelas dari data tersebut.

Klasifikasi terdiri dari dua proses yaitu:

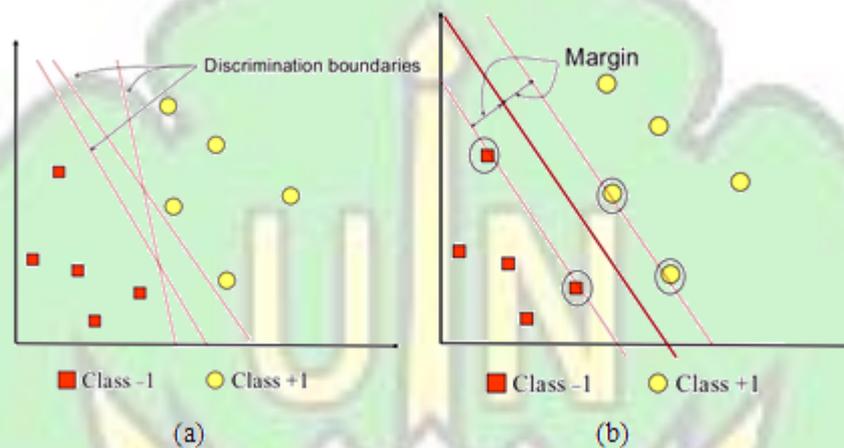
1. *Process training* merupakan suatu proses untuk mencari pemodelan dengan menggunakan *data training*. Pada penelitian ini peneliti menggunakan algoritma SVM. Pada SVM ini dilakukan dengan memberi label pada masing-masing nilai fitur.
2. *Process testing* merupakan suatu proses untuk menguji suatu pemodelan yang sudah didapatkan dari data training. Penggunaan model untuk mengklasifikasi data baru. Di sini, catatan data disimpan dalam model dan model memberikan respons kelas dari hasil perhitungan.

D. SVM

SVM pertama kali diperkenalkan pada tahun 1992 oleh *Vapnik* dengan *Partner Boser* dan *Guyon*. Prinsip dasar SVM adalah *classifier linear*, yang kemudian dikembangkan untuk memecahkan masalah *non-linear* dengan mengintegrasikan konsep *trik kernel* ke dalam area kerja dimensi tinggi[21]. SVM dapat mengklasifikasikan data linier dan non linier. Variabel-variabel prediktor merupakan data input sedangkan variabel target yang saling bergantung merupakan output. SVM bertujuan untuk menemukan fungsi klasifikasi terbaik dan untuk membedakan antara anggota dari dua kelas dalam *data training*. Matrik untuk konsep fungsi klasifikasi "terbaik" dapat diwujudkan secara geometris. Untuk dataset terpisah secara linear, fungsi

klasifikasi linier berhubungan dengan *hyperplane* pemisah $f(x)$ yang melewati tengah dua kelas, memisahkan keduanya[17].

Model algoritma SVM merupakan salah satu algoritma dari metode klasifikasi, yang bekerja dengan cara mencari suatu garis (*hyperplane*) untuk memisahkan dua kelompok data. Berikut ini adalah contoh berdasarkan pada Gambar 1 tentang bagaimana SVM mencoba menemukan *hyperplane* terbaik untuk memisahkan kelas -1 dan +1:



Gambar 2. 2 SVM mencoba menemukan *hyperplane* terbaik untuk memisahkan kedua class -1 dan +1[25].

Gambar 2.2 Menunjukkan beberapa pola yang merupakan anggota dari dua kelas, yaitu -1 dan +1. Pola di kelas -1 dilambangkan dengan warna merah (kotak), sedangkan pola di kelas +1 dilambangkan dengan warna kuning (bulat). Masalah klasifikasi dapat diselesaikan dengan mencoba mencari garis (*hyperplane*) yang memisahkan dua kelas. Gambar 1-a menunjukkan batas *diferensiasi alternatif* yang berbeda. *Hyperplane* dengan pemisah terbaik dapat ditemukan dengan mengukur margin dari *hyperplane* dan mencari titik maksimum. Kernel harus digunakan untuk mencapai keberhasilan banyak algoritma klasifikasi untuk permukaan linier [25]. Dengan demikian dapat diketahui bahwa jenis kernel dapat mempengaruhi hasil klasifikasi yang dilakukan.

Hyperplane adalah garis pemisah terbaik antara dua kelas. Untuk mencari *hyperplane* dapat dilakukan dengan mencari *margin hyperplane*

dan mencari titik maksimum. Margin adalah jarak antara data terdekat di antara dua kelas yang berbeda, yang disebut dengan support vektor[26]. Garis solid pada gambar 1-b menunjukkan *hyperplane* yang terbaik, karena terletak tepat diantara kedua class, sedangkan *support vector* dilambangkan dengan titik merah dan kuning yang berada di dalam lingkaran hitam.

Hyperplane klasifikasi *linear SVM* dinotasikan:

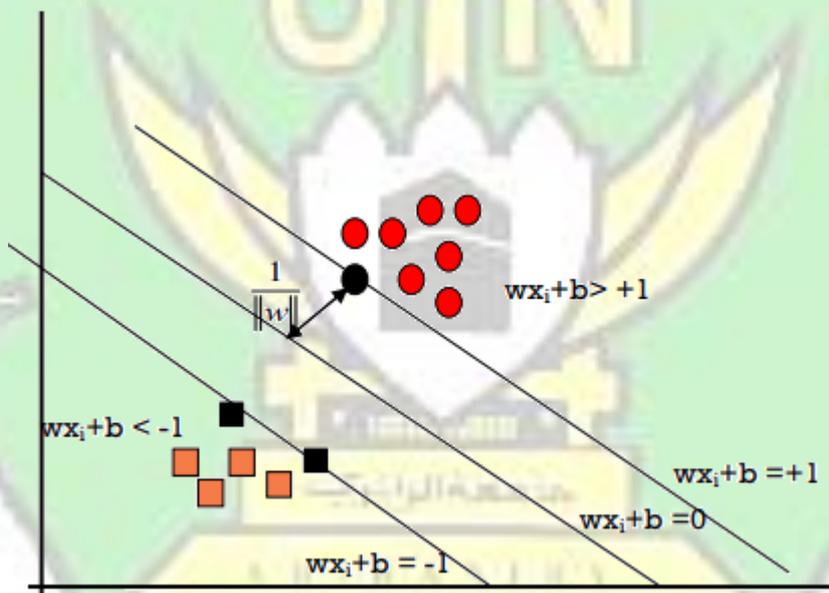
$$f(x) = w \cdot x + b = 0 \quad (1)$$

Dari persamaan di atas di dapatkan pertidaksamaan kelas +1 (negatif)

$$w \cdot x + b \leq +1 \quad (2)$$

Pertidaksamaan kelas -1:

$$w \cdot x + b \geq -1 \quad (3)$$



Gambar 2. 3 Temukan fungsi pemisahan optimal untuk objek yang dapat dipisahkan secara linear [27].

w adalah bidang normal dan b adalah posisi bidang relatif terhadap koordinat pusat. Dengan mengoptimalkan nilai jarak antara *hyperplane* dan titik berikutnya, margin terbesar dapat ditemukan, yaitu $1 / \|w\|$. Ini dapat dirumuskan sebagai masalah pemrograman kuadratik (QP) di mana

titik minimum persamaan (4) dengan mengingat kendala dari persamaan tersebut(5).

$$\min \frac{1}{2} \|w\|^2 = \min \frac{1}{2}(w_1^2 + w_2^2) \quad (4)$$

$$y_i (w_{xi} + b) \geq 1, \quad i = 1, 2, 3 \dots, N(5)$$

1. Kelebihan dan kekurangan SVM

Kelebihan-kelebihan *SVM* sebagai berikut[28]:

- a. *Generalisasi* dapat diartikan sebagai kemampuan suatu metode *SVM* untuk mengklasifikasikan pola yang tidak berisi data yang digunakan dalam fase pembelajaran metode ini.
- b. *Curse Of Dimensionality* adalah masalah yang biasanya dihadapi ketika proses pengenalan pola ketika memperkirakan parameter. Karena jumlah sampel data relatif kecil dibandingkan dengan ruang data vektor, sehingga semakin tinggi ruang vektor yang diproses, ini mengarah pada konsekuensi yang memerlukan jumlah data dengan tiga dimensi[29].
- c. *Feasibility*
SVM dapat diimplementasikan dengan sangat mudah karena proses penentuan support vektor dalam masalah QP dapat dirumuskan. Jadi jika kita memiliki perpustakaan untuk menyelesaikan masalah QP, *SVM* itu sendiri dapat digunakan dengan sangat mudah[25].

Adapun kekurangan *SVM* sebagai berikut[25]:

- a. Sulit digunakan untuk pengolahan data yang mempunyai jumlah data yang besar.
- b. Metode *SVM* secara teoritis dikembangkan untuk masalah klasifikasi dengan dua atau lebih kelas. Namun masing-masing strategi ini memiliki kelemahan, sehingga untuk pengembangan

SVM pada masalah yang lebih dari dua kelas masih menjadi topik penelitian terbuka[28].

E. Naive Bayes

Naive Bayes adalah perhitungan statistik yang menghitung kemungkinan yang serupa antara kasus lama dengan kasus baru. *Naive Bayes* memiliki kecepatan yang baik ketika diterapkan ke database besar dan mempunyai tingkat akurasi yang tinggi[30]. *Naive Bayes* adalah bagian dari pembelajaran yang diawasi, sehingga proses pembelajaran membutuhkan data awal dalam bentuk data pelatihan untuk membuat keputusan. Nilai probabilitas akan dihitung ketika proses klasifikasi dari masing-masing label class yang di input. Label class dengan nilai probabilitas tertinggi digunakan sebagai label class untuk data input. *Naive Bayes* adalah perhitungan teori *Bayesian* yang paling sederhana karena dapat mengurangi kompleksitas komputasi dalam penggantian probabilitas yang sederhana. Selain itu, algoritma *Naive Bayes* juga dapat memproses data yang memiliki banyak atribut[22].

F. Bayes Network

Bayesian network adalah model grafis probabilistik sederhana (PGM), yang didasarkan pada teori *probabilitas* dan teori grafik. Teori probabilitas secara langsung terkait dengan data, sedangkan teori grafik secara langsung terkait dengan bentuk representasi yang ingin dipertahankan. Misalnya, *Bayesian network* dapat mewakili hubungan probabilistik antara penyakit dan gejala. *Bayesian network* berfungsi untuk menghitung kemungkinan penyebab dari berbagai gejala penyakit[31].

G. Decision Tree

Decision tree adalah diagram alur yang hampir sesuai dengan struktur pohon. Setiap internal node menunjuk atribut yang akan diuji, masing-masing cabang mempresentasikan hasil dari atribut tes tersebut, dan *leaf node*

mempresentasikan kelas tertentu atau distribusi kelas. *Decision tree* ada 3 jenis node, yaitu[19]:

1. *Root node* adalah simpul teratas. Tidak ada input pada node ini dan bisa tidak memiliki output memiliki lebih dari satu output.
2. *Internal node* adalah simpul cabang. Terdapat satu input dan mempunyai output minimal dua pada node ini
3. *Leaf node* atau terminal node adalah node akhir, hanya terdapat satu input dan tidak mempunyai output pada node ini.

H. Weka

Weka adalah sebuah paket *tools machine learning* praktis, dan juga sebuah *software open source* yang di keluarkan di bawah GNU *general public license* dan berbasis java. “Weka” merupakan singkatan dari “*Waikato Environment For Knowledge Analysis*”, yang dibuat di *universitas Waikato New Zealand* untuk penelitian, pendidikan dan beberapa aplikasi. Weka dapat menyelesaikan masalah klasifikasi pada data mining di dunia nyata. *Software* ini dibuat menggunakan bahasa java yang dapat berjalan hampir di semua *platform*. Weka berisi alat untuk data *preprocessing, classify, cluster, associate, select attribute* dan *visualize*. Weka merupakan aplikasi data mining tanpa berbayar yang berbasis Java, yang terdiri dari beberapa algoritma yang digunakan untuk proses generalisasi dan formulasi dari kumpulan data[32].

I. Metode Evaluasi

a. *Confusion Matrix*

Confusion matrix merupakan alat ukur dalam bentuk matriks 2x2, yang digunakan untuk mendapatkan akurasi dari klasifikasi pada setiap data kelas lulus dan tidak lulus. Setiap kelas yang diprediksi memiliki empat kemungkinan output yang berbeda, yaitu true positif (TP) dan true negative (TN), yang menunjukkan klasifikasi yang benar[20]. Jika output yang diprediksi *positif* sedangkan nilai asli *negatif* itu disebut *false positive* (FP), dan jika output yang diprediksi *negatif* sedangkan nilai asli *positif*,

itu disebut *false negative* (FN). Berikut ini dalam Tabel 1 disajikan bentuk confusion matriks seperti yang dijelaskan sebelumnya.

Tabel 2. 1 Model Confusion matrix untuk masalah klasifikasi kelas yang diprediksi[20]

		Nilai Sebenarnya	
		<i>TRUE</i>	<i>FALSE</i>
Nilai Prediksi	<i>TRUE</i>	<i>True Positive</i> (<i>TP</i>)	<i>False Negatif</i> (<i>FP</i>)
	<i>FALSE</i>	<i>False Positive</i> (<i>FN</i>)	<i>True Negative</i> (<i>TN</i>)

- *True positives* adalah jumlah *record positif* yang diklasifikasikan sebagai positif.
- *False positives* adalah jumlah *record negatif* yang diklasifikasikan sebagai positif.
- *False negatives* adalah jumlah *record positif* yang diklasifikasikan sebagai negatif.
- *True negatives* adalah jumlah *record negatif* yang diklasifikasikan sebagai negatif.

Untuk menghitung nilai *precision*, *recall*, *accuracy* dan *F-Measure* dapat dilakukan dengan memasukkan data training ke dalam *confusion matrix*[20].

- *Precision* merupakan tingkat kesesuaian antara output dari sistem dengan informasi yang diperlukan oleh user.
- *Recall* merupakan tingkat keberhasilan sistem dalam mencari informasi.
- *Accuracy* diartikan sebagai tingkat keseimbangan antara nilai aktual dari nilai estimasi.
- *F-Measure* adalah perhitungan yang menggabungkan nilai *recall* dan nilai *precision*. Nilai *recall* dan *Precision* dalam situasi tertentu memiliki prediksi yang berbeda. Ukuran yang

menunjukkan timbal balik antara *Recall* dan *Precision* adalah *F-Measure* adalah bobot rata-rata dari harmonik dan recall dan *precision*.

Nilai *precision* dan *recall* di dapatkan dari (9) dan (10), sedangkan untuk akurasi dan *F-Measure precision* dari (11) dan (12)[33].

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (8)$$

$$F - Measure = 2 \times \frac{precision \times recall}{precision + recall} \quad (9)$$

Keterangan :

TP = *True Positives*, TN = *True Negatives*, FP = *False Positives*, FN = *False Negatives*.

J. Penelitian Terdahulu

Penelitian terdahulu menjadi paduan penulis dalam melakukan penelitian. Penulis akan memaparkan beberapa penelitian terdahulu dengan tujuan untuk menjelaskan perbedaan dalam penelitian yang dilakukan oleh penulis dan membandingkannya dengan penelitian sebelumnya. Dengan demikian penelitian yang dilakukan oleh penulis terbukti asli. Berikut ini adalah penelitian yang berkaitan dengan penelitian yang dilakukan oleh penulis.

Penelitian yang pertama yang dilakukan oleh Sri Widaningsih mahasiswa prodi teknik informatika tahun 2019 dengan judul “*Memprediksi tingkat*

kelulusan mahasiswa teknik informatika dengan menggunakan algoritma C4.5, Naïve Bayes dan KNN,”. Penelitian ini bertujuan untuk memprediksi tingkat kelulusan mahasiswa dan IPK standar menggunakan metode data mining dengan fungsi klasifikasi. Metodologi yang digunakan dalam fase data mining ini adalah *Discovery Knowledge Database (KDD)* dimulai dari tahap seleksi, pra-pemrosesan, transformasi, data mining dan evaluasi. Ada tiga algoritma yang digunakan dalam penelitian ini yaitu *C4.5, k-nearest neighbor (kNN,)* dan *Naïve Bayes*. Hasil akhir dari ketiga algoritma menunjukkan bahwa algoritma *Naïve Bayes* adalah algoritma terbaik untuk memprediksi kelulusan mahasiswa tepat waktu dengan $IPK \geq 3$ dengan akurasi (76,79%), kesalahan (23,17%) dan AUC (0,850)[17].

Penelitian yang kedua yang dilakukan oleh Yuandri Trisaputra pada tahun 2016 tentang “*Klasifikasi Profil Siswa SMA/SMK yang Masuk PTN (Perguruan Tinggi Negeri) dengan k-Nearest Neighbor*”. Penelitian ini bertujuan untuk memprediksi kampus pilihan yang menerima siswa dengan model terbaik yang diperoleh. Menggunakan algoritma *k-Nearest Neighbor* sebagai sebagai pengklasifikasi untuk membuat model klasifikasi untuk data siswa yang diterima oleh Kampus Pilihan Mahasiswa di SNMPTN. Hasil akhirnya, metode *k-Nearest Neighbor* dapat digunakan untuk mengklasifikasikan data SNMPTN dengan akurasi yang diperoleh dari *classifier* terbaik adalah 83.3607 % [23].

Selanjutnya penelitian terdahulu yang dilakukan oleh Syarli dengan judul “*Memprediksi kelulusan mahasiswa menggunakan metode naive bayes (Studi Kasus: Data Mahasiswa Baru Universitas Al Asyariah Mandar sulawesi Barat)*”. penelitian ini bertujuan untuk melakukan prediksi peluang kelulusan mahasiswa baru di perguruan tinggi. Hasil akhir yaitu *Naïve Bayes* dapat memprediksi peluang masa depan berdasarkan pengalaman masa lalu dengan nilai akurasi yang akurat menunjukkan keefektifan dataset Penerimaan Mahasiswa Baru yang diterapkan ke dalam metode *Naïve Bayes Classification*. Implementasi *Naïve Bayes* dengan aplikasi WEKA dapat melacak karakteristik atribut dari dataset dengan opsi pilihan lulus.

Pengelompokan pilihan lulus berdasarkan atribut yang dipilih, yaitu program studi, pilihan pertama, pilihan kedua dan rata-rata[10].



BAB III

METODOLOGI PENELITIAN

A. Spesifikasi Perangkat

Proses klasifikasi dokumen web dalam penelitian ini menggunakan laptop dengan *Spesifikasi Processor AMD E1-1200 APU with Radeon (TM) HD Graphics 1.40 GHz, sistem operasi Windows 7 Ultimate, RAM 2 GB.*

B. Tempat dan Waktu Penelitian

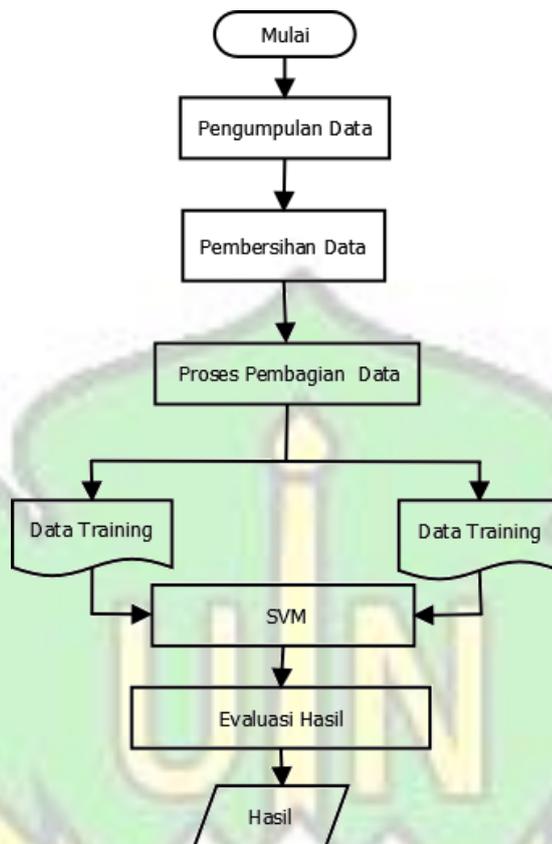
Penelitian ini mulai dilakukan pada bulan Agustus 2019. Data didapatkan melalui kuesioner yang dibagikan kepada mahasiswa baru tahun ajaran 2019 yang ada di beberapa universitas Banda Aceh.

C. Teknik Pengumpulan Data

Pengumpulan data merupakan langkah yang sangat penting dalam metode penelitian[34]. Penelitian ini dilakukan dengan membagi kuesioner kepada mahasiswa baru tahun ajaran 2019 yang ada di beberapa universitas Banda Aceh, dengan menggunakan teknik *simple random sampling* yaitu pengambilan sampel yang dilakukan secara acak pada seluruh populasi tanpa memperhatikan tingkatan yang ada didalamnya, Teknik ini memberikan kesempatan yang sama pada seluruh populasi untuk dijadikan sebagai sampel data.

D. Prosedur Penelitian

Metode yang digunakan dalam penelitian ini yaitu model algoritma SVM. Prosedur penelitian yang akan dilakukan pada penelitian ini dapat dilihat pada gambar 3.1 :



Gambar 3. 1 Prosedur penelitian

Keterangan dari prosedur penelitian di atas adalah:

a. Pengumpulan Data

1) Studi pustaka

Peneliti mengumpulkan berbagai referensi dan literatur pendukung penelitian yang berkaitan dengan penggunaan *SVM* dalam memprediksi kelulusan jalur masuk perguruan tinggi, literatur dapat berupa buku teks, jurnal dan artikel dan karya ilmiah lainnya.

2) Kuesioner

Kuesioner adalah teknik pengumpulan data yang menggunakan serangkaian pertanyaan tertulis yang diberikan kepada responden untuk dijawab[5]. Dalam penelitian ini kuesioner dibuat menggunakan *tool google form* yang dibagikan baik secara

langsung ataupun melalui sosial media. Kuesioner terdiri dari 22 variabel serta jawaban yang sudah disediakan oleh peneliti untuk dipilih oleh responden. Variabel dari kuesioner dapat dilihat pada tabel 3.1 Berikut ini.

Tabel 3. 1 Variabel kuesioner

No	Variabel	No	Variabel
1	Jumlah saudara kandung	12	Prestasi di sekolah
2	Ortu yang dimiliki sekarang	13	Les privat di sekolah
3	Siapa yang membiayai sekolah	14	Berapa lama les privat
4	Pendidikan terakhir ayah	15	Bimbel seleksi masuk PTN
5	Pendidikan terakhir ibu	16	Berapa lama bimbel
6	Pekerjaan ayah	17	Nilai UN
7	Pekerjaan ibu	18	Jalur yang diikuti
8	Asal sekolah	19	Jalur minat
9	Jurusan di sekolah	20	Jalur lulus
10	Organisasi di sekolah	21	Sesuai atau tidak dengan jalur minat
11	Organisasi di luar sekolah	22	Jurusan lulus

b. Proses Pembersihan Data

Data yang diperoleh dari hasil pembagian kuesioner sebanyak 826 data, data tersebut dapat dilihat pada (Gambar 3.2) di bawah ini. Namun tidak semua data dan atribut dapat digunakan karena harus melalui beberapa tahap pemrosesan data awal. Tahap awal dari *pre processing* adalah proses pembersihan data. Untuk mendapatkan data berkualitas, beberapa tahapan dapat dilakukan selama pemrosesan data, yaitu[35]:

a. *Data validation*

biru adalah variabel yang sudah melalui proses cleaning , kolom yang berwarna putih adalah variabel korelasi sedangkan tabel yang berwarna hijau dan kuning adalah variabel yang digunakan untuk labelnya. Warna hijau untuk label korelasi sedangkan yang warna kuning untuk label implementasi di weka.

c. Transformasi data

Transformasi data merupakan proses pengubahan data ke dalam format atau bentuk *Arff* sesuai dengan format perangkat lunak *WEKA*. Data yang di dapatkan melalui microsoft excel diubah ke dalam format CSV “,” atau data yang menggunakan tanda pemisah dengan koma (dapat dilihat pada gambar 3.2) selanjutnya diubah kedalam format *Arff* melalui *WEKA*(dapat dilihat pada gambar 3.3).

```
File Edit Format View Help
Apa pendidikan terakhir ayah anda?,Apa pendidikan terakhir ibu anda?,Apa pekerjaan ayah anda?,Apa pekerjaan ibu ar
PERGURUAN TINGGI,PERGURUAN TINGGI,Pegawai Negeri,Ibu Rumah Tangga,Tidak,Tidak,Tidak,99,SNMPTN,UM-PTKIN,TIDAK SESUAI
SMA/SEDERAJAT,PERGURUAN TINGGI,Tidak Bekerja,Pegawai Negeri,Ya,Ya,Tidak,46,SNMPTN,PMB,TIDAK SESUAI
PERGURUAN TINGGI,SMA/SEDERAJAT,Buruh,Ibu Rumah Tangga,Tidak,Tidak,Tidak,76,SNMPTN,PMB,TIDAK SESUAI
PERGURUAN TINGGI,SMA/SEDERAJAT,Tukang,Ibu Rumah Tangga,Tidak,Ya,Ya,45,UM-PTKIN,PMB,TIDAK SESUAI
PERGURUAN TINGGI,PERGURUAN TINGGI,Buruh,Buruh,Tidak,Tidak,Tidak,52,SNMPTN,PMB,TIDAK SESUAI
SMA/SEDERAJAT,SMA/SEDERAJAT,Tidak Bekerja,Pegawai Negeri,Ya,Ya,Tidak,50,SNMPTN,SBMPTN,TIDAK SESUAI
SMA/SEDERAJAT,SMP/SEDERAJAT,Pegawai Negeri,Pegawai Negeri,Tidak,Ya,Tidak,41,UM-PTKIN,PMB,TIDAK SESUAI
PERGURUAN TINGGI,PERGURUAN TINGGI,wiraswasta,Ibu Rumah Tangga,Tidak,Ya,Tidak,40,SBMPTN,SBMPTN,SESUAI
SMA/SEDERAJAT,PERGURUAN TINGGI,Tidak Bekerja,Ibu Rumah Tangga,Tidak,Ya,Tidak,85,SNMPTN,UM-PTKIN,TIDAK SESUAI
SMA/SEDERAJAT,SMA/SEDERAJAT,Buruh,Ibu Rumah Tangga,Tidak,Tidak,Tidak,48,SNMPTN,PMB,TIDAK SESUAI
SMA/SEDERAJAT,PERGURUAN TINGGI,Pegawai Negeri,Ibu Rumah Tangga,Tidak,Ya,Tidak,65,SNMPTN,PMB,TIDAK SESUAI
PERGURUAN TINGGI,PERGURUAN TINGGI,wiraswasta,Ibu Rumah Tangga,Tidak,Ya,Ya,87,SNMPTN,SBMPTN,TIDAK SESUAI
SMA/SEDERAJAT,SMA/SEDERAJAT,wiraswasta,Pegawai Negeri,Tidak,Ya,Tidak,56,SBMPTN,UM-PTKIN,TIDAK SESUAI
TIDAK SEKOLAH,SMA/SEDERAJAT,Pegawai Negeri,Pegawai Negeri,Tidak,Tidak,Tidak,33,SNMPTN,SNMPTN,SESUAI
SMA/SEDERAJAT,SMA/SEDERAJAT,Pegawai Negeri,Pegawai Negeri,Tidak,Ya,Ya,55,SBMPTN,SBMPTN,SESUAI
PERGURUAN TINGGI,SMA/SEDERAJAT,wiraswasta,Pegawai Negeri,Ya,Ya,Tidak,76,PMB,PMB,SESUAI
PERGURUAN TINGGI,SMA/SEDERAJAT,Buruh,Ibu Rumah Tangga,Tidak,Ya,Tidak,50,SPAN-PTKIN,PMB,TIDAK SESUAI
TIDAK SEKOLAH,SMP/SEDERAJAT,wiraswasta,Ibu Rumah Tangga,Tidak,Ya,Tidak,55,SBMPTN,PMB,TIDAK SESUAI
SMA/SEDERAJAT,SMA/SEDERAJAT,Buruh,Ibu Rumah Tangga,Ya,Tidak,Tidak,90,SPAN-PTKIN,PMB,TIDAK SESUAI
SMA/SEDERAJAT,PERGURUAN TINGGI,wiraswasta,Ibu Rumah Tangga,Tidak,Ya,Tidak,51,SBMPTN,PMB,TIDAK SESUAI
SD/SEDERAJAT,SD/SEDERAJAT,Pegawai Negeri,Ibu Rumah Tangga,Tidak,Tidak,Tidak,54,SNMPTN,SBMPTN,TIDAK SESUAI
```

Gambar 3. 2 Sampel data dalam format CSV

```

@relation 'DATA BARU 810'

@attribute 'Berapa Jumlah saudara kandung anda?' numeric
@attribute 'Orang tua yang anda miliki sekarang?' {'Ayah &
ibu','Ibu saja ( ayah sudah meninggal)','ibu & ayah sudah
meninggal','Ayah saja ( ibu sudah meninggal)','Ayah dan ibu','Ayah
& ibu, Ayah dan ibu tetapi sudah cerai','Cerai','Tinggal sama
ayah','Ayah & ibu, Ibu dan ayah bercerai',Ibu}
@attribute 'Siapa yang membiayai sekolah anda?' {'Bapak saja','Ibu
saja','Bapak & ibu','Abang dan ibu','Orang lain/wali','Bapak & ibu,
Ibu saja','biaya sendiri','Bapak & ibu, biaya sendiri','nenek','Abang
Ipar','Bapak & ibu, Orang lain/wali',Nenek,Kakak}
@attribute 'Apa pendidikan terakhir ayah anda?' {'PERGURUAN
TINGGI',SMA/SEDERAJAT,'TIDAK SEKOLAH',SD/SEDERAJAT,SMP/SEDERAJAT}
@attribute 'Apa pendidikan terakhir ibu anda?' {'PERGURUAN
TINGGI',SMA/SEDERAJAT,SMP/SEDERAJAT,SD/SEDERAJAT,'TIDAK SEKOLAH'}
@attribute 'Apa pekerjaan ayah anda?' {'Pegawai BUMN','Sakit (Tidak
Bekerja)',Petani,Sopir,'Sudah meninggal','Pegawai
negeri','Wiraswasta,BUMN,Tani','Tni al','Buruh Kasar','Sudah
mninggal','Pedagang biasa',Polisi,'Jual gas keliling',Buruh,'Tidak

```

Gambar 3. 3 Sampel data dalam format Arff

Setelah melalui proses pembersihan data, data akhir yang didapatkan dalam penelitian ini berjumlah 809 data, kemudian data tersebut dibagi menjadi dua yaitu data training dan data testing. Sebelumnya pada penelitian ini terdapat beberapa variabel (dapat dilihat pada tabel 3.1).

Selanjutnya untuk mengetahui variabel yang berpengaruh antara variabel-variabel tersebut maka harus dilakukan proses korelasi dengan menggunakan *toll SPSS*. Proses pengolahan data menggunakan *SPSS* harus menggunakan data numerik, maka seluruh data dari variabel-variabel tersebut harus dilakukan *convert* ke data berbentuk numerik, data hasil korelasi dapat dilihat gambar sebelumnya pada (Tabel 3.3).

c. Proses Pembagian Data

Setelah melalui cleaning proses selanjutnya data di analisis dan dikelompokkan antar variabel yang berhubungan satu sama lainnya. Setelah data dianalisis kemudian diterapkan model-model yang sesuai dengan jenis data. Untuk membuat pemodelan perlu dilakukan pembagian data yaitu pembagian data training dan data testing dimana data training

dijadikan sebagai model dan data testing untuk melihat keakuratan sebuah model.

Pada penelitian ini untuk proses pembagian data menggunakan model pengujian. Model pengujian merupakan alat yang digunakan untuk melakukan pengujian dalam penelitian ini[11]. Adapun alat untuk pengujian pada weka tersedia berbagai macam, seperti *use training set*, *supplied test set*, *cross validation* dan *percentage split*. Namun pada penelitian ini, jenis pengujian yang digunakan adalah *cross validation* dan *percentage split* [11]. Berikut penjelasannya.

1) *Cross-Validation*

Pada *cross-validation*, akan ada inputan user untuk memasukkan nilai *fold* yang digunakan. Di dalam weka nilai defaultnya adalah 10. Pada penelitian ini, nilai *fold* yang digunakan adalah 5 dan 10. Berikut penjelasan mengenai nilai *fold* 5 dan 10.

- Nilai *Fold* 5

Nilai *fold* 5 artinya adalah terjadi 5 lipatan yang berulang-ulang, lalu masing-masing bagian secara bergantian melakukan pengujian. Dimana pada sekali perulangan yang menjadi data testing sebanyak 1 kali dan yang menjadi *data training* sebanyak k-1 kali yaitu 4. Setelah melakukan *cross validation* 5 kali lipatan dan menghitung hasil evaluasi, weka memanggil algoritma pembelajaran pada waktu terakhir kali (ke-6) pada seluruh dataset untuk mendapatkan model yang diuji kembali hingga selesai.

- Nilai *Fold* 10

Nilai *fold* 10 artinya terjadi 10 lipatan pada dataset yang terus berulang-ulang. sehingga 1 menjadi data testing dan k-1 menjadi *data training*.

Jadi pada penelitian ini dengan menggunakan model pengujian *cross-validation* peneliti akan menggunakan nilai *k-fold* 5 dan *k-fold* 10 sebagai data pengujian.

2) *Percentage Split*

Percentage split merupakan hasil testing dengan menggunakan $k\%$ dari data yang dimasukkan. Dimana nilai k adalah input dari user. Pada penelitian ini, *percentage split* yang digunakan adalah 20%, 30%, 40%, dan 50% dengan hasil uji pengukuran evaluasi yang digunakan adalah *precision*, *recall* dan *f-measure*.

Berikut adalah penjelasan *split* 20%, 30%, 40% dan 50% :

- *Split* 20% artinya 20% dari total data menjadi data pelatihan dan sisanya yaitu 80% menjadi data tes.
- *Split* 30% artinya 30% dari total data menjadi data pelatihan dan sisanya yaitu 70% menjadi data tes.
- *Split* 40% artinya 40% dari total data menjadi data pelatihan dan sisanya yaitu 60% menjadi data tes.
- *Split* 50% artinya 50% dari total data menjadi data pelatihan dan sisanya yaitu 50% menjadi data tes.

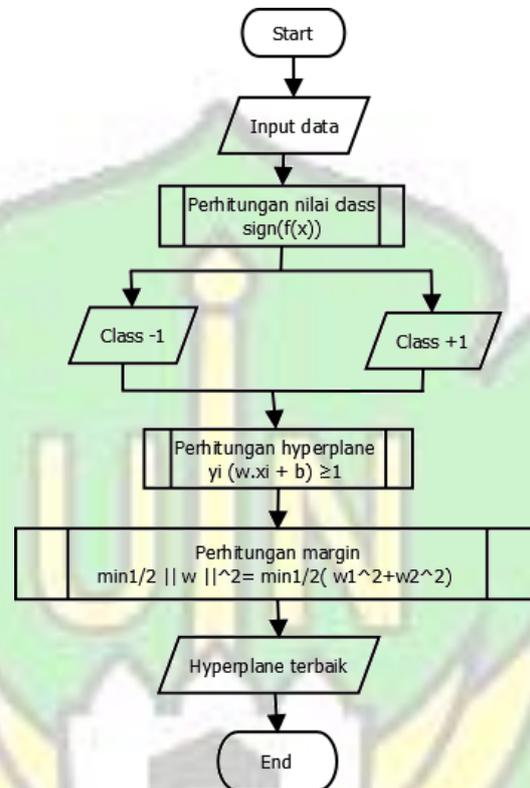
Jadi pada penelitian ini dengan menggunakan model pengujian *percentage split* peneliti akan menggunakan nilai *split* 20%, 30%, 40% dan 50% sebagai data pengujian.

d. SVM

SVM merupakan salah satu algoritma dari metode klasifikasi, yang bekerja dengan cara mencari suatu garis (*hyperplane*) terbaik untuk memisahkan dua kelompok data. Berikut tahapan proses kerja SVM:

Tahapan pertama yaitu menginput data, kemudian data tersebut akan dicari nilai kelasnya yaitu kelas +1(negatif) dan kelas -1(positif) dengan menggunakan rumus persamaan (1), setelah diketahui kelasnya kemudian akan dicari dua data dari kelas yang berbeda yang memiliki jarak terdekat (*support vector*) selanjutnya dari dua data tersebut dicari garis pemisah (*hyperplane*) dengan menggunakan rumus persamaan (2), dan terakhir di cari jarak maksimum dari

hyperplane dengan *support vector* (*margin*) dengan menggunakan persamaan 5 untuk mengetahui *hyperplane* terbaik dari dua data tersebut. Berikut flowchart proses kerja SVM:



Gambar 3. 4 Flowchart SVM

Contoh studi kasus SVM[38]:

Tabel 3. 4 Contoh studi kasus SVM

X_1	X_2	Kelas (y)
1	1	1
1	-1	-1
-1	1	-1
-1	-1	-1

- Karena ada dua fitur (x_1 dan x_2), maka w juga memiliki dua fitur (w_1 dan w_2).
- Formulasi yang digunakan adalah sebagai berikut:

$$\min \frac{1}{2} || w ||^2$$

$$y_i (w_{xi} + b) \geq 1, i = 1,2,3 \dots, N$$

$$y_i (w_1 \cdot x_1 + w_2 \cdot x_2 + b) \geq 1$$

Mencari persamaan:

$$y_1 (w_1 \cdot x_1 + w_2 \cdot x_2 + b) \geq 1$$

$$1 (w_1 \cdot 1 + w_2 \cdot 1 + b) \geq 1$$

$$w_1 + w_2 + b \geq 1 \quad \text{Persamaan 1}$$

$$y_2 (w_1 \cdot x_1 + w_2 \cdot x_2 + b) \geq 1$$

$$-1 (w_1 \cdot 1 + w_2 \cdot (-1) + b) \geq 1$$

$$-1(w_1 - w_2 + b) \geq 1$$

$$(-w_1 + w_2 - b) \geq 1 \quad \text{Persamaan 2}$$

$$y_3 (w_1 \cdot x_1 + w_2 \cdot x_2 + b) \geq 1$$

$$-1 (w_1 \cdot (-1) + w_2 \cdot 1 + b) \geq 1$$

$$-1(-w_1 + w_2 + b) \geq 1$$

$$(w_1 - w_2 - b) \geq 1 \quad \text{Persamaan 3}$$

$$Y_4 (w_1 \cdot x_1 + w_2 \cdot x_2 + b) \geq 1$$

$$-1 (w_1 \cdot (-1) + w_2 \cdot (-1) + b) \geq 1$$

$$-1(-w_1 - w_2 + b) \geq 1$$

$$(w_1 + w_2 - b) \geq 1 \quad \text{Persamaan 4}$$

Didapatkan beberapa persamaan berikut:

a. $(w_1 + w_2 + b) \geq 1$ untuk $y_1= 1, x_1=1, x_2=1$

b. $(-w_1 + w_2 - b) \geq 1$ untuk $y_2= -1, x_1=1, x_2= -1$

c. $(w_1 - w_2 - b) \geq 1$ untuk $y_3= -1, x_1=1, x_2= -1$

d. $(w_1 + w_2 - b) \geq 1$ untuk $y_4= -1, x_1=1, x_2= -1$

➤ Menjumlahkan persamaan (1) dan (2) :

$$(w_1 + w_2 + b) \geq 1$$

$$(-w_1 + w_2 - b) \geq 1$$

----- +

$$2w_2 = 2$$

$$w_2 = 1$$

➤ Menjumlahkan persamaan (1) dan (3):

$$(w_1 + w_2 + b) \geq 1$$

$$(w_1 - w_2 - b) \geq 1$$

----- +

$$2w_1 = 2$$

$$w_1 = 1$$

➤ Menjumlahkan persamaan (2) dan (3) :

$$(-w_1 + w_2 - b) \geq 1$$

$$(w_1 - w_2 - b) \geq 1$$

----- +

$$-2b = 2$$

$$b = -1$$

Sehingga di dapatkan persamaan hyperplane : Dengan margin :

$$w_1 \cdot x_1 + w_2 \cdot x_2 + b = 0$$

$$\min \frac{1}{2} \|w\|^2 = \min \frac{1}{2} (w_1^2 + w_2^2)$$

$$1 \cdot x_1 + 1 \cdot x_2 - 1 = 0$$

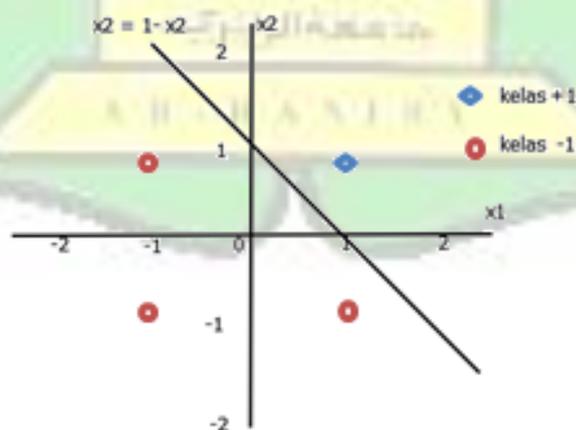
$$\frac{1}{2} \|w\|^2 = \frac{1}{2} (1^2 + 1^2)$$

$$x_1 + x_2 - 1 = 0$$

$$\frac{1}{2} \|w\|^2 = 2$$

$$x_2 = 1 - x_1$$

$$\|w\|^2 = 2$$



Gambar 3. 5 Visualisasi garis hyperplane

e. Evaluasi Hasil

Berdasarkan hasil pengujian dari model yang telah ditetapkan diatas maka akan didapatkan beberapa hasil akurasi dari *confusion matrix*. Dari *confusion matrix* dapat dicari akurasi dari *precision*, *recall*, *accuracy* dan *f-measure*, sebagai contoh dapat dilihat pada penyelesaian di bawah ini dengan menggunakan rumus (6), (7), (8) & (9).

Tabel 3. 5 Contoh confusion matrik

Classification	Predicted Class	
Nilai Prediksi	True Negatif = 451	False Positif = 49
	False Negatif = 127	True Positif = 141

Evaluasi hasil dari confusion matrix pada tabel 3.2 sebagai berikut:

$$\begin{aligned}
 Precision &= \frac{TP}{TP + FP} \\
 &= \frac{141}{141 + 49} \\
 &= \frac{141}{190} = 0.742 = 74 \%
 \end{aligned}$$

$$\begin{aligned}
 Recall &= \frac{TP}{TP + FN} \\
 &= \frac{141}{141 + 127} \\
 &= \frac{141}{268} = 0.526 = 52 \%
 \end{aligned}$$

$$\begin{aligned}
 Accuracy &= \frac{TP + TN}{TP + TN + FN + FP} \\
 &= \frac{141 + 451}{141 + 451 + 127 + 49} \\
 &= \frac{592}{768} = 0.770 = 77 \%
 \end{aligned}$$

$$\begin{aligned}
 F - \text{Measure} &= 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \\
 &= 2 \times \frac{0.742 \times 0.526}{0.742 + 0.526} \\
 &= 2 \times \frac{0.390}{1.268} \\
 &= \frac{0.78}{1.268} = 0.615 = 61 \%
 \end{aligned}$$

e. Hasil

Setelah proses klasifikasi yang dilakukan di atas maka hasil akhir yang akan didapatkan dari penelitian ini yaitu berupa akurasi.



BAB IV HASIL DAN PEMBAHASAN

A. Dataset

Pada penelitian ini peneliti menggunakan data mahasiswa baru tahun ajaran 2019 dari beberapa perguruan tinggi di Banda Aceh. Pengumpulan data dilakukan melalui kuesioner, data yang digunakan mengenai kelulusan jalur masuk perguruan tinggi, yang terdiri dari 809 data.

1. Labeling

Labeling merupakan proses penentuan label pada data. Pada penelitian ini “jalur lulus” menjadi labelnya. Karena jalur lulus merupakan hasil akhir dari pemilihan jalur masuk perguruan tinggi. Kemudian diberikan kelas pada label tersebut dengan melakukan convert data yang dimana variabel jalur lulus awalnya berbentuk kategori, kemudian diubah menjadi numerik.

Tabel 4. 1 Pemberian nama kelas pada label

JALUR LULUS	NUMERIK	JUMLAH JALUR LULUS
SNMPTN	1	54
SBMPTN	2	126
UM-PTKIN	3	138
SPAN-PTKIN	4	57
PMB	5	433

B. Hasil Penelitian

1. Uji Korelasi

Untuk mengetahui hubungan antar variabel (H_1 atau H_0) maka dapat dilihat berdasarkan nilai *signifikansi*, nilai *r hitung* dan arah nilai *pearson correlation*. Untuk ketentuan nilainya dapat di lihat pda tabel 4.2 berikut ini[37] :

Tabel 4. 2 Hubungan korelasi

Hubungan Korelasi	Signifikasi	r- hitung(Pearson Correlation)	Arah Nilai Pearson Correlation	
			(+)	(-)
H_1	< 0.05	$> r\text{-tabel}$	Hubungan searah	Hubungan tidak searah
H_0	> 0.05	$< r\text{-tabel}$		

Keterangan :

- H1= Terdapat hubungan antar variabel yang dihubungkan
- H0 = Tidak terdapat hubungan antar variabel yang dihubungkan
- Hubungan searah = Jika variabel X bernilai tinggi maka variabel Y juga bernilai tinggi
- Hubungan tidak searah = Jika variabel X bernilai tinggi maka variabel Y bernilai rendah dan begitu sebaliknya.

Pada penelitian ini peneliti hendak menguji apakah terdapat hubungan antara variabel yang terliha pada tabel 3.3 terhadap jalur lurus. Dalam hal ini peneliti menjadikan jalur lurus sebagai label akhir untuk menentukan kelulusan mahasiswa di jalur masuk perguruan tinggi.

Dalam menetapkan interval kategori kekuatan korelasi penetapannya sebagai berikut :

Tabel 4. 3 Kategori Kekuatan Korelasi

0	Tidak ada korelasi
0,00-0,25	Korelasi sangat lemah
0,25-0,50	Korelasi cukup
0,50-0,75	Korelasi kuat
0,75-0,99	Korelasi sangat kuat
1	Korelasi sempurna

Dari kategori di atas dapat disimpulkan bahwa jika hubungan suatu variabel tidak sama dengan nol (0), maka dapat dikatakan terjadi hubungan pada variabel tersebut. Berikut hasil korelasi antar variabel dalam penelitian ini:

		Jalur minat	Jalur lurus
Jalur minat	Pearson Correlation	1	.311**
	Sig. (2-tailed)		.000
	N	809	809
Jalur lurus	Pearson Correlation	.311**	1
	Sig. (2-tailed)	.000	
	N	809	809

** . Correlation is significant at the 0.01 level (2-tailed).

Gambar 4. 1 Korelasi jalur minat dengan jalur lurus

Dari gambar di atas dapat dilihat bahwa korelasi antara jalur minat berhubungan positif dengan jalur lurus dengan nilai pearson

correlation sebesar 0,311** yang memiliki arti bahwa korelasi cukup, dikarenakan nilai pearson correlation lebih besar dari r tabel 0,065 (0,311 > 0,065), kemudian nilai signifikansi sebesar 0,000 dengan demikian pada variabel ini terdapat hubungan yang signifikan dikarenakan nilai signifikansi lebih kecil dari 0,05 (0,000 < 0,05) sehingga dapat dikatakan variabel di atas menolak H0. Dengan nilai determinasi $0,311^2 = 0,096721$ atau 10 %.

		Bimbel seleksi masuk PTN	Jalur lulus
Bimbel seleksi masuk PTN	Pearson Correlation	1	-.108**
	Sig. (2-tailed)		.002
	N	809	809
Jalur lulus	Pearson Correlation	-.108**	1
	Sig. (2-tailed)	.002	
	N	809	809

** . Correlation is significant at the 0.01 level (2-tailed).

Gambar 4. 2 Korelasi bimbel seleksi masuk PTN dengan jalur lulus

Dari gambar di atas terlihat bahwa korelasi antara dua variabel tersebut berhubungan negatif dengan nilai Pearson Correlation dari korelasi variabel “bimbel seleksi masuk PTN terhadap jalur minat” sebesar -0,108 dengan r tabel 0,065, maka dapat disimpulkan bahwa memiliki korelasi sangat lemah antara variabel yang dihubungkan, kemudian nilai signifikansi sebesar 0,002 artinya terdapat hubungan yang signifikan antara variabel yang dihubungkan karena nilai signifikansi lebih kecil dari 0,05 (0,002 < 0,05) sehingga dapat dikatakan variabel di atas menolak H0. Dengan nilai determinasi $-0,108^2 = 0,011664$ atau 1 %.

		Nilai UN	Jalur lulus
Nilai UN	Pearson Correlation	1	.014
	Sig. (2-tailed)		.684
	N	809	809
Jalur lulus	Pearson Correlation	.014	1
	Sig. (2-tailed)	.684	
	N	809	809

Gambar 4. 3 Korelasi nilai UN dengan jalur lulus

Dari gambar di atas dapat dilihat bahwa korelasi antara nilai UN berhubungan positif dengan jalur lulus dengan nilai *pearson correlation*

sebesar 0,014 yang memiliki arti bahwa korelasi sangat lemah, dikarenakan nilai *pearson correlation* lebih kecil dari r tabel 0,065 ($0,014 < 0,065$), kemudian nilai signifikansi sebesar 0,684 dengan demikian pada variabel ini tidak terdapat hubungan yang signifikan dikarenakan nilai signifikansi lebih besar dari 0,05 ($0,684 > 0,05$) sehingga dapat dikatakan variabel di atas menolak H1. Dengan nilai determinasi $0,014^2=0,000196$ atau 0 %.

		Organisasi di SMA	Jalur lulus
Organisasi di SMA	Pearson Correlation	1	.036
	Sig. (2-tailed)		.304
	N	809	809
Jalur lulus	Pearson Correlation	.036	1
	Sig. (2-tailed)	.304	
	N	809	809

Gambar 4. 4 Korelasi organisasi di SMA dengan jalur lulus

Dari gambar di atas dapat dilihat bahwa korelasi antara organisasi di SMA berhubungan positif dengan jalur lulus dengan nilai *pearson correlation* sebesar 0,036 yang memiliki arti bahwa korelasi sangat lemah, dikarenakan nilai *pearson correlation* lebih kecil dari r tabel 0,065 ($0,311 > 0,065$), kemudian nilai signifikansi sebesar 0,304 yang dapat disimpulkan bahwa tidak terdapat hubungan yang signifikan pada variabel tersebut, karena nilai signifikansi lebih besar dari 0,05 ($0,304 > 0,05$) sehingga dapat dikatakan variabel di atas menolak H1. Dengan nilai determinasi $0,036^2=0,001296$ atau 0 %.

		Les private di SMA	Jalur lulus
Les private di SMA	Pearson Correlation	1	-.018
	Sig. (2-tailed)		.607
	N	809	809
Jalur lulus	Pearson Correlation	-.018	1
	Sig. (2-tailed)	.607	
	N	809	809

Gambar 4. 5 Korelasi les private di SMA dengan jalur lulus

Dari gambar di atas terlihat bahwa korelasi antara dua variabel tersebut berhubungan negatif dengan nilai *Pearson Correlation* sebesar -0,018 dengan r tabel 0,065, maka dapat disimpulkan bahwa memiliki

korelasi sangat lemah antara variabel yang dihubungkan, kemudian nilai signifikansi sebesar 0,607 artinya tidak terdapat hubungan antar variabel yang dihubungkan yang signifikan karena nilai signifikansi lebih besar dari 0,05 (0,60 > 0,05) sehingga dapat dikatakan variabel di atas menolak H1. Dengan nilai determinasi $-0,018^2 = 0,000324$ atau 0 %.

		Pekerjaan ayah	Jalur lulus
Pekerjaan ayah	Pearson Correlation	1	.086*
	Sig. (2-tailed)		.015
	N	809	809
Jalur lulus	Pearson Correlation	.086*	1
	Sig. (2-tailed)	.015	
	N	809	809

*. Correlation is significant at the 0.05 level (2-tailed).

Gambar 4. 6 Korelasi pekerjaan ayah dengan jalur lulus

Dari gambar di atas dapat dilihat bahwa korelasi antara pekerjaan ayah berhubungan positif dengan jalur lulus dengan nilai *pearson correlation* sebesar 0,086 yang memiliki arti bahwa korelasi sangat lemah, dikarenakan nilai *pearson correlation* lebih kecil dari *r* tabel 0,065 (0,086 < 0,065), kemudian nilai signifikansi sebesar 0,015 yang dapat disimpulkan bahwa terdapat hubungan yang signifikan antara pekerjaan ayah dan jalur lulus, karena nilai signifikansi lebih kecil dari 0,05 (0,015 < 0,05) sehingga dapat dikatakan variabel di atas menolak H0. Dengan nilai determinasi $0,086^2 = 0,007396$ atau 1%.

		Pekerjaan ibu	Jalur lulus
Pekerjaan ibu	Pearson Correlation	1	.045
	Sig. (2-tailed)		.203
	N	809	809
Jalur lulus	Pearson Correlation	.045	1
	Sig. (2-tailed)	.203	
	N	809	809

Gambar 4. 7 Korelasi pekerjaan ibu dengan jalur lulus

Dari gambar di atas dapat dilihat bahwa korelasi antara pekerjaan ibu berhubungan positif dengan jalur lulus dengan nilai *pearson correlation* sebesar 0,045 yang memiliki arti bahwa korelasi sangat lemah, dikarenakan nilai *pearson correlation* lebih kecil dari *r* tabel 0,065

($0,045 < 0,065$), kemudian nilai signifikansi sebesar 0,203 yang dapat disimpulkan bahwa tidak terdapat hubungan yang signifikan antara pekerjaan ibu dan jalur lulus, karena nilai signifikansi lebih besar dari 0,05 ($0,203 > 0,05$) sehingga dapat dikatakan variabel di atas menolak H1. Dengan nilai determinasi $0,045^2 = 0,002025$ atau 0 %.

		Pendidikan terakhir ayah	Jalur lulus
Pendidikan terakhir ayah	Pearson Correlation	1	.024
	Sig. (2-tailed)		.497
	N	809	809
Jalur lulus	Pearson Correlation	.024	1
	Sig. (2-tailed)	.497	
	N	809	809

Gambar 4. 8 Korelasi pendidikan terakhir ayah dengan jalur lulus

Dari gambar di atas dapat dilihat bahwa korelasi antara pendidikan terakhir ayah berhubungan positif dengan jalur lulus dengan nilai *pearson correlation* sebesar 0,024 yang memiliki arti bahwa korelasi sangat lemah, dikarenakan nilai *pearson correlation* lebih kecil dari r tabel 0,065 ($0,024 < 0,065$), kemudian nilai signifikansi sebesar 0,497 yang dapat disimpulkan bahwa tidak terdapat hubungan yang signifikan antara pendidikan terakhir ayah dan jalur lulus, karena nilai signifikansi lebih besar dari 0,05 ($0,497 > 0,05$) sehingga dapat dikatakan variabel di atas menolak H1. Dengan nilai determinasi $0,024^2 = 0,001764$ atau 0,18%.

		Pendidikan terakhir ibu	Jalur lulus
Pendidikan terakhir ibu	Pearson Correlation	1	-.015
	Sig. (2-tailed)		.667
	N	809	809
Jalur lulus	Pearson Correlation	-.015	1
	Sig. (2-tailed)	.667	
	N	809	809

Gambar 4. 9 Korelasi pendidikan terakhir ibu dengan jalur lulus

Dari gambar di atas terlihat bahwa korelasi antara variabel tersebut memiliki hubungan yang negatif dengan nilai *pearson correlation* sebesar -0,015 dengan r tabel 0,065, maka dapat disimpulkan bahwa memiliki korelasi sangat lemah antara variabel yang dihubungkan, kemudian nilai

signifikansi sebesar 0,667 artinya tidak terdapat hubungan yang signifikan antara variabel yang dihubungkan karena nilai signifikansi lebih besar dari 0,05 ($0,667 > 0,05$) sehingga dapat dikatakan variabel di atas menolak H1. Dengan nilai determinasi $-0,015^2 = 0,000225$ atau 0 %.

		Orang tua yang dimiliki sekarang	Jalur lulus
Orang tua yang dimiliki sekarang	Pearson Correlation	1	.049
	Sig. (2-tailed)		.165
	N	809	809
Jalur lulus	Pearson Correlation	.049	1
	Sig. (2-tailed)	.165	
	N	809	809

Gambar 4. 10 Korelasi orang tua yang dimiliki dengan jalur lulus

Dari gambar di atas dapat dilihat bahwa korelasi antara orang tua yang dimiliki sekarang berhubungan positif dengan jalur lulus dengan nilai *pearson correlation* sebesar 0,049 yang memiliki arti bahwa korelasi sangat lemah, dikarenakan nilai *pearson correlation* lebih kecil dari r tabel 0,065 ($0,049 > 0,065$), kemudian nilai signifikansi sebesar 0,165 yang dapat disimpulkan bahwa tidak terdapat hubungan yang signifikan antara orang tua yang dimiliki dan jalur lulus, karena nilai signifikansi lebih besar dari 0,05 ($0,165 > 0,05$) sehingga dapat dikatakan variabel di atas menolak H1. Dengan nilai determinasi $0,049^2$ kuadrat = 0,002401 atau 0 %.

Tabel 4. 4 Perbandingan Korelasi Variabel

Variabel	Sig	Pearson Correlation	Keterangan
Jalur minat	0,000	0,311**	Korelasi Cukup
Bimbel seleksi masuk PTN	0,002	-0,180**	Korelasi Sangat lemah
Nilai UN	0,684	-0,014	Korelasi Sangat Lemah
Organisasi di SMA	0,304	0,036	Korelasi Sangat Lemah
Les private di SMA	0,607	-0,018	Korelasi Sangat Lemah
Pekerjaan ayah	0,015	0,086*	Korelasi Sangat Lemah
Pekerjaan ibu	0,203	0,045	Korelasi Sangat Lemah

Pendidikan terakhir ayah	0,497	0,024	Korelasi Sangat Lemah
Pendidikan terakhir ibu	0,667	-0,015	Korelasi Sangat Lemah
Orang tua yang dimiliki sekarang	0,165	0,049	Korelasi Sangat Lemah

Setelah melakukan perhitungan diatas dapat diketahui bahwa variabel yang memiliki hubungan positif (artinya variabel X bernilai tinggi dan variabel Y juga bernilai tinggi) dengan variabel jalur lulus yaitu jalur minat, organisasi di SMA, pekerjaan ayah, pekerjaan ibu, pendidikan terakhir ayah dan orang tua yang dimiliki sekarang, sedangkan yang berhubungan negatif (artinya variabel X bernilai tinggi maka variabel Y bernilai rendah) dengan variabel jalur lulus yaitu bimbel seleksi masuk PTN, les private di SMA dan pendidikan terakhir ibu.

Dari kesimpulan di atas dapat disimpulkan bahwa terdapat dua variabel yang memiliki hubungan korelasi yang lebih baik daripada variabel lain yaitu variabel jalur minat dan variabel bimbel seleksi masuk PTN, jalur minat memiliki korelasi sebesar 0,311** ($r=0,311$) dan bimbel seleksi masuk memiliki korelasi sebesar -0,180** ($r= 0,180$).

1. Pengujian Klasifikasi

Pada pengujian ini, peneliti menggunakan aplikasi weka untuk menguji keakuratan dari algoritma SVM.

a. Perbandingan Uji Variabel

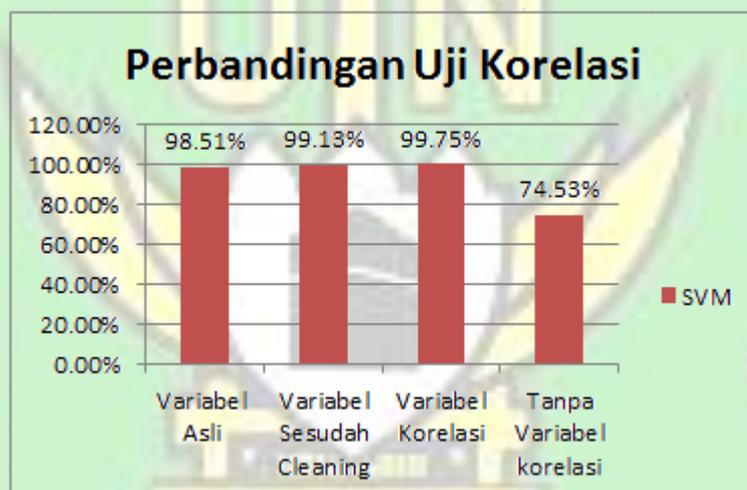
Pada penelitian ini peneliti ingin menguji hasil akurasi antara variabel asli yang berjumlah 22 variabel dengan variabel yang sudah di cleaning yang berjumlah 11 variabel , variabel korelasi yang berjumlah 2 variabel dan tanpa variabel korelasi berjumlah 7 variabel.

Tabel 4. 5 Perbanding uji variabel

Pengujian	Variabel			
	Variabel Asli	Variabel Sesudah Cleaning	Variabel Korelasi	Tanpa Variabel korelasi
SVM	98.51 %	99.13 %	99.75%	74.53%

Berdasarkan tabel di atas dapat diketahui bahwa dengan menggunakan algoritma SVM, variabel asli memiliki akurasi sebesar

98.51%, variabel sesudah cleaning memiliki akurasi sebesar 99.13% , variabel korelasi memiliki akurasi 99.75% dan tanpa variabel korelasi memiliki akurasi 74.53%. Jadi dapat disimpulkan bahwa nilai akurasi variabel asli dengan variabel yang sudah di cleaning tidak jauh berbeda karena di kedua variabel tersebut sama-sama terdapat variabel yang sangat berhubungan dengan variabel class, sedangkan antara variabel berkorelasi dengan tanpa variabel korelasi memiliki perbedaan akurasi yang sangat signifikan , dikarenakan pada variabel korelasi tersebut hanya terdapat variabel yang sangat berpengaruh dengan classnya, sedangkan pada tanpa variabel korelasi sama sekali tidak terdapat variabel yang berhubungan dengan variabel class. Berikut adalah grafik dari tabel di atas:



Grafik 4. 1 Perbandingan akurasi pada variabel asli, variabel sesudah cleaning, variabel korelasi dan tanpa variabel korelasi

Dari tabel di atas dapat disimpulkan bahwa variabel korelasi memiliki nilai akurasi yang lebih tinggi dibandingkan dengan variabel lainnya, karena variabel tersebut memiliki hubungan korelasi yang baik dengan variabel class label.

2. Jenis Pengujian

Pada *Weka* terdapat beberapa metode pengujian yaitu *Use training set*, *Supplied test set*, *Percentage split* dan *Cross-Validation*. Akan tetapi dalam

penelitian ini peneliti menggunakan dua metode pengujian saja yaitu *cross-validation* dan *percentage split* untuk mendapatkan hasil dari metode SVM [11].

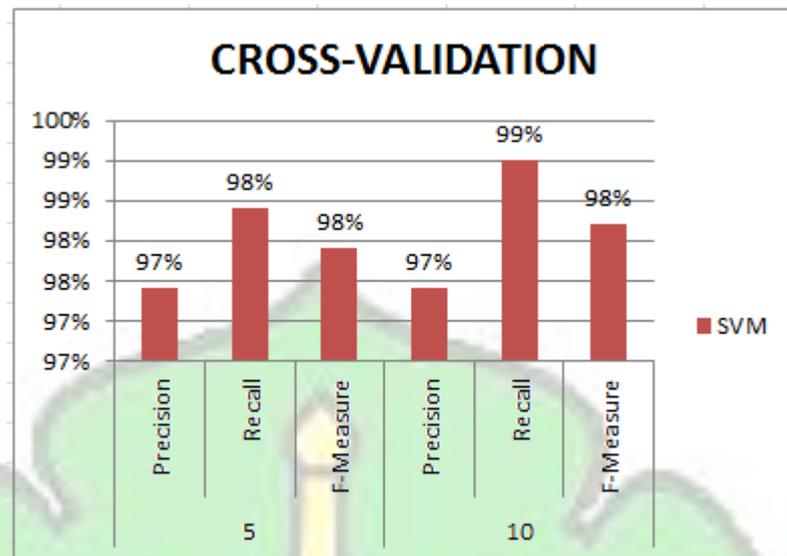
a. Cross-validation

Pada pengujian ini nilai *fold* yang digunakan adalah *5-fold* dan *10-fold* untuk mendapatkan nilai akurasi dari *precision*, *recall*, dan *f-measure*. Berikut adalah hasil yang diperoleh:

Tabel 4. 6 Cross Validation

Cross Validation	Akurasi	SVM
5	Precision	97%
	Recall	98%
	F-Measure	98%
10	Precision	97%
	Recall	99%
	F-Measure	98%

Berdasarkan tabel di atas dapat diketahui bahwa pada metode SVM, dengan menggunakan *fold 5* memiliki akurasi yaitu *precision 97%*, *recall 98%* dan *f-measure 98%* dan dengan menggunakan *Fold 10* memiliki akurasi yaitu *precision 97%*, *recall 99%* dan *f-measure 98%*. Dari hasil akurasi *precision*, *recall* dan *f-measure* pada metode SVM dapat dilihat bahwa tidak terdapat perbedaan antara nilai *k-fold 5* dan *k-fold 10*. Berikut ini adalah grafik hasil *cross-validation* berdasarkan tabel di atas:



Grafik 4. 2 Cross Validation

Berdasarkan grafik di atas dapat diketahui bahwa tidak terdapat perbedaan antara nilai *precision*, *recall* dan *f-measure* yang diterapkan pada metode SVM. Jadi dapat disimpulkan bahwa nilai *k-fold* pada *cross-validation* tidak mempengaruhi hasil akurasi klasifikasi pada metode SVM.

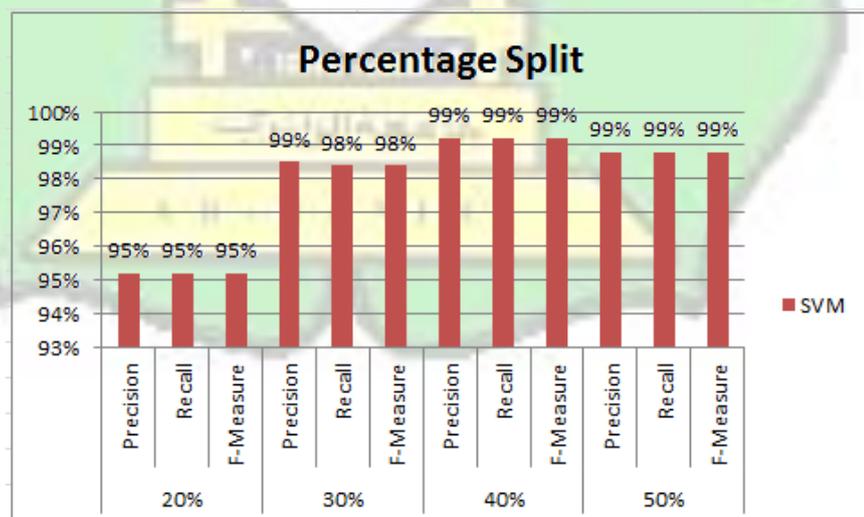
b. Percentage Split

Percentage split adalah jumlah data training yang akan diuji. pada penelitian ini 4 nilai *split* yang digunakan diantaranya yaitu 20%,30%,40% dan 50%. Berikut ini pada tabel 4.6 hasil yang ditampilkan dari klasifikasi menggunakan metode SVM .

Tabel 4. 7 Percentage Split

Percentage Split	Akurasi	SVM
20%	Precision	95%
	Recall	95%
	F-Measure	95%
30%	Precision	99%
	Recall	98%
	F-Measure	98%
40%	Precision	99%
	Recall	99%
	F-Measure	99%
50%	Precision	99%
	Recall	99%
	F-Measure	99%

Berdasarkan tabel diatas dapat diketahui bahwa klasifikasi percentage split pada metode SVM dengan nilai split terendah adalah pada split 20% memiliki akurasi 95% dan dengan split tertinggi yaitu pada split 50% memiliki akurasi 99%. Dari hasil akurasi percentage split tersebut dapat disimpulkan bahwa semakin tinggi nilai split yang diberikan maka akan semakin tinggi nilai akurasinya. Berikut ini adalah grafik hasil percentage split berdasarkan tabel diatas:

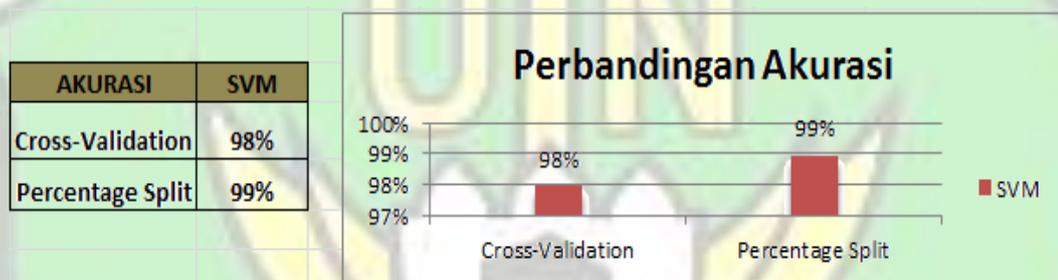


Grafik 4. 3 Percentage Split

Dari grafik diatas dapat disimpulkan bahwa split dapat mempengaruhi hasil akurasi yang diperoleh dari metode SVM, dan split yang paling baik digunakan oleh metode SVM di antara keempat nilai split itu adalah split 50%, jadi semakin tinggi nilai *split* maka akan semakin tinggi nilai akurasinya.

c. Perbandingan Jenis Pengujian

Berdasarkan hasil tertinggi yang diperoleh dari hasil klasifikasi menggunakan metode SVM dengan pengujian *cross-validation* dan *percentage split*, maka selanjutnya dilakukan perbandingan untuk mendapatkan hasil terbaik. Berikut pada gambar 4.11 adalah hasil perbandingan nilai akurasi antara *cross-validation* dan *percentage split*:

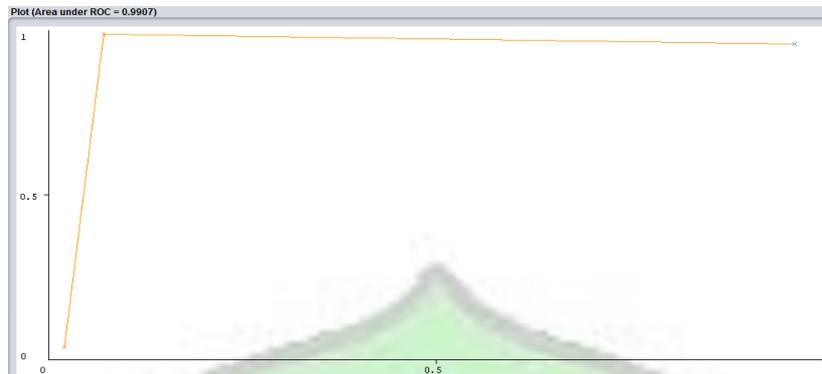


Gambar 4. 11 Perbandingan Jenis Pengujian

Berdasarkan tabel di atas dapat diketahui bahwa pada jenis pengujian *cross-validation* memiliki nilai akurasi 98% sedangkan pada pengujian *percentage split* memiliki nilai akurasi 99%. Jadi dari perbandingan di atas dapat disimpulkan bahwa *percentage split* memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan *cross-validation* yaitu dengan nilai akurasi 99%.

d. Kurva ROC

Hasil perhitungan yang divisualisasi dengan kurva ROC, dapat dilihat pada gambar 4.12 di bawah ini yang merupakan kurva dengan algoritma SVM.



Gambar 4. 12 Kurva ROC dengan Algoritma SVM

Dari gambar di atas dapat dilihat kurva ROC memiliki nilai AUC sebesar 0.9907 yang dapat dikategorikan sebagai klasifikasi yang sangat baik.

4. Analisis Hasil

Model yang dihasilkan dengan metode SVM yang diuji menggunakan metode *cross validation* [6], terlihat perbandingan nilai accuracy, precision, recall dan *f-measure* pada tabel 4.7 berikut ini:

Tabel 4. 8 Accuracy dan AUC

SVM	
Accuracy	99%
AUC	0.9907

Tabel di atas membandingkan accuracy dan AUC dari tiap metode. Terlihat bahwa nilai accuracy menunjukkan nilai yang sesuai. Untuk klasifikasi data mining, nilai AUC dapat dibagi menjadi beberapa kelompok[36]:

1. 0.90-1.00 = klasifikasi sangat baik
2. 0.80-0.90 = klasifikasi baik
3. 0.70-0.80 = klasifikasi cukup

4. 0.60-0.70 = klasifikasi buruk
5. 0.50-0.60 = klasifikasi salah

Berdasarkan pengelompokan di atas maka dapat disimpulkan bahwa algoritma SVM termasuk klasifikasi sangat baik pada proses klasifikasi dalam “Memprediksi Kelulusan Jalur Masuk Perguruan Tinggi” karena memiliki nilai AUC antara 0.90-1.00 dan memiliki akurasi sebesar 99%.

a. Perbandingan Nilai Akurasi Dengan Algoritma Lain

Pada tabel 4.9 dapat dilihat perbandingan nilai akurasi. Algoritma SVM memiliki akurasi 99% sedangkan algoritma Naive Bayes memiliki akurasi 87%, dari perbandingan tersebut dapat diketahui bahwa algoritma SVM memiliki tingkat akurasi terbaik.

Tabel 4. 9 Perbandingan akurasi

SVM	NAIVE BAYES
99%	87%

5. Waktu Konsumsi

Waktu konsumsi adalah waktu yang diperlukan untuk membangun model pada suatu percobaan[36]. Waktu konsumsi terhitung dalam satuan *second*. Untuk mengukur waktu konsumsi, maka digunakan salah satu pembahasan sebelumnya untuk mengukur konsumsi waktu yang diperlukan. Berikut ini merupakan perbandingan waktu konsumsi antara pengujian *percentage split* dengan *cross validation*.

Tabel 4. 10 Waktu konsumsi

Percentage Split	Waktu Konsumsi	Cross Validation	Waktu konsumsi
20%	0.02	5 Fold	0.5
30%	0.02		
40%	0.02	10 Fold	0.5
50%	0.01		

Dari tabel di atas dengan menggunakan dua pengujian yaitu *percentage split* dan *cross validation* dapat diketahui bahwa pada *percentage split* 20% (artinya 20% sebagai data training yang akan memprediksi sisanya yaitu 80% sebagai data testing) dengan waktu yang dibutuhkan sebesar 0.02 *second*. Kemudian *split* 30% membutuhkan waktu sebanyak 0.02 *second*, *split* 40% membutuhkan waktu sebanyak 0.02 *second* dan pada *split* 50% membutuhkan waktu lebih sedikit dibandingkan yang lain yaitu sebesar 0.01 *second*. Selanjutnya dengan menggunakan pengujian Cross validation dengan *k-fold* 5 membutuhkan waktu sebanyak 0.5 *second* dan dengan menggunakan *k-fold* 10 membutuhkan waktu sebanyak 0.5 *second*. Dari hasil waktu konsumsi antara dua pengujian tersebut dapat disimpulkan pengujian *percentage split* sangat cocok digunakan dalam memprediksi dengan algoritma SVM karena dapat memprediksi dengan waktu yang sangat cepat.

BAB V

KESIMPULAN DAN SARAN

A. KESIMPULAN

1. Dari berbagai cara yang dipersiapkan oleh calon mahasiswa hanya variabel bimbel yang bisa menjamin calon mahasiswa untuk lulus di jalur yang diminati dengan nilai korelasi sebesar $-0,180^{**}$ dan nilai signifikansi sebesar $0,002$.
2. Dari hasil implementasi algoritma *SVM* dalam memprediksi kelulusan calon mahasiswa di setiap seleksi jalur masuk perguruan tinggi berdasarkan hasil evaluasi dan validasi dapat disimpulkan bahwa algoritma *SVM* memiliki akurasi dan kinerja sangat baik yaitu rata-rata sebesar 99% dengan nilai AUC (Area Under Curve) sebesar 0.9907 dan waktu konsumsi yang dibutuhkan hanya sebesar $0.01-0.02$ *second*.

B. SARAN

Saran yang dapat diberikan untuk pengembangan penelitian selanjutnya, sebagai berikut:

1. Dalam penelitian ini hanya menggunakan aplikasi machine learning weka sebagai alat bantu, diharapkan pada penelitian selanjutnya, dapat melakukan penelitian menggunakan aplikasi tambahan untuk mendukung dan memperkuat hasil akurasi dari klasifikasi dokumen.
2. Untuk pengembangan selanjutnya bisa melakukan perbandingan dengan berbagai algoritma, dan menggunakan alat pengujian dengan nilai yang lebih tinggi.
3. Bagi peneliti yang ingin melanjutkan penelitian ini agar bisa menambahkan variabel baru ataupun mengganti variabel lain yang memiliki tingkat korelasi yang lebih tinggi dengan variabel labelnya.

DAFTAR PUSTAKA

- [1] P. D. S. 2008 Abbas, "MANAGEMEN PERGURUAN TINGGI 2008." .
- [2] F. Abdullah, "TINGGI AGAMA ISLAM NEGERI DI SULAWESI SELATAN Fahri Abdullah Institut Agama Islam Negeri Bone , Indonesia," vol. 17, pp. 683–693, 2018.
- [3] R. fajar Salinggih, "IMPLEMENTASI NAIVE BAYES UNTUK MEMPREDIKSI CALON MAHASISWA BARU YANG AKAN LANJUT MELAKUKAN PEMBAYARAN ATAU TIDAK (Studi Kasus : PMB UIN Bandung)," p. 2019, 2019.
- [4] B. Jalur, M. Snmptn, S. D. A. N. Mandiri, U. N. Makassar, D. Lestari, and H. Anis, "ANALISIS HASIL BELAJAR MAHASISWA JURUSAN FISIKA THE ANALYSIS OF LEARNING OUTCOMES OF STUDENTS DEPARTMENT OF PHYSICS BASED ON ENTRANCE (SNMPTN , SBMPTN , AND MANDIRI) UNIVERSITAS NEGERI MAKASSAR," pp. 15–25, 2018.
- [5] file:///C:/Users/Marhamah/Downloads/9044-21027-1-S. pdfJajan. J. Purnama, N. Ichsan, E. Ermawati, T. Wahyuni, and A. Suryadi, "Analisis Pengaruh Kesiapan Belajar Terhadap Hasil Simulasi SBMPTN 2018 (Studi Kasus Peserta Simulasi SBMPTN 2018)," vol. 14, no. 2, pp. 14–23, 2018.
- [6] <https://span-ptkin.ac.id/>, "INFORMASI SPAN-PTKIN," p. 2014, 2014.
- [7] U. N. Makassar, "Perbedaan kecerdasan emosional dan hasil belajar Kalkulus Mahasiswa Jurusan Kimia ditinjau dari jalur masuk," pp. 472–474, 2015.
- [8] P. A. I. F. T. K. Uinsa and U. M. Ptkin, "TARBIYAH DAN KEGURUAN UIN SUNAN AMPEL SURABAYA (Analisis Perbandingan Penerimaan Jalur SPAN , UM PTKIN dan Jalur Mandiri Tahun 2016)," pp. 244–263, 2017.
- [9] <https://uin-ar-raniry.ac.id/index.php/id/posts/uin-ar-raniry-kembali-buka-jalur-pmb-lokal>, "uin-ar-raniry-kembali-buka-jalur-pmb-lokal," p. 24,

2018.

- [10] S. Syarli and A. A. Muin, "Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi)," *J. Ilmu Komput.*, vol. 2, no. 1, pp. 22–26, 2016.
- [11] R. SITANGGANG, "Analisis Kinerja Metode Naïve Bayes dan SSVM untuk Menentukan Pola Kelompok Penyakit," 2017.
- [12] S. M. Andri, Yesi Novaria Kunang, "Implementasi Teknik Data Mining Untuk Memprediksi Tingkat Kelulusan," vol. 2013, no. June 2016, pp. 56–63, 2013.
- [13] S. Kasus, P. Mahasiswa, B. Politeknik, R. Hidayat, and M. Ula, "Penentuan Kelulusan Calon Mahasiswa Jalur SNMPTN Menggunakan Fuzzy Inference System Mamdani."
- [14] F. Munawaroh, "PENGARUH JALUR MASUK TERHADAP PRESTASI MAHASISWA PROGRAM," no. May, 2018.
- [15] <https://um-ptkin.ac.id/home/informasi>, "INFORMASI UM-PTKIN," p. 2019, 2019.
- [16] D. Teori, "PENGUNAAN METODE NAÏVE BAYES CLASSIFIER PADA APLIKASI PERPUSTAKAAN."
- [17] D. A. C, N. Bayes, and D. A. N. Svm, "PERBANDINGAN METODE DATA MINING UNTUK PREDIKSI NILAI DAN WAKTU KELULUSAN MAHASISWA PRODI TEKNIK INFORMATIKA," vol. 13, no. 1, pp. 16–25, 2019.
- [18] D. S. Kusumo, M. A. Bijaksana, and D. Darmantoro, "Data Mining Dengan Algoritma Apriori Pada Rdbms Oracle," *TEKTRIKA - J. Penelit. dan Pengemb. Telekomun. Kendali, Komputer, Elektr. dan Elektron.*, vol. 8, no. 1, pp. 1–5, 2019.
- [19] Y. N. Kunang, Andri, and S. Murniati, "Implementasi Teknik Data Mining Untuk Memprediksi Tingkat," vol. 2013, no. A-56-A-63, pp. 1–8, 2013.
- [20] S. Defiyanti and M. Kom, "Analisis dan Prediksi Kinerja Mahasiswa Menggunakan Teknik Data Mining," vol. 2, pp. 1–8, 2013.
- [21] R. U. S. U. Sitanggang, "Analisis Kinerja Metode Naïve Bayes dan SSVM

- untuk Menentukan Pola Kelompok Penyakit,” 2017.
- [22] D. Sartika, D. I. Sensuse, U. Indo, G. Mandiri, and F. I. Komputer, “Perbandingan Algoritma Klasifikasi Naive Bayes , Nearest Neighbour , dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian,” vol. 1, no. 2, pp. 151–161, 2017.
- [23] Y. Trisaputra, “Klasifikasi Profil Siswa SMA / SMK yang Masuk PTN (Perguruan Tinggi Negeri) dengan k-Nearest Neighbor Klasifikasi Profil Siswa SMA / SMK yang Masuk PTN (Perguruan Tinggi Negeri) dengan k-Nearest Neighbor Yuandri Trisaputra , Indriyani , Shellafuri Mar,” no. September 2015, 2016.
- [24] M. A. Banjarsari, H. I. Budiman, and A. Farmadi, “Penerapan K-Optimal Pada Algoritma Knn untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan IP Sampai Dengan Semester 4,” vol. 02, no. 02, pp. 50–64, 2015.
- [25] A. S. Nugroho, A. B. Witarto, and D. Handoko, “SVM,” 2003.
- [26] R. A. Permana and S. Sahara, “Metode SVM Sebagai Penentu Kelulusan Mahasiswa pada Pembelajaran Elektronik,” vol. VII, no. 1, pp. 50–58, 2019.
- [27] B. Santosa and S. Surabaya, “1 . Ide Dasar SVM,” no. x, 1995.
- [28] I. B. dan I. S. S. Kurniawan, N. Z. (2016). TA: Identifikasi Jenis Penyakit Daun Tembakau Menggunakan Metode Gray Level Co-Occurrence Matrix (GLCM) dan SVM (SVM) (Doctoral dissertation, “No Title,” pp. 6–33.
- [29] D. E. Ratnawati and A. W. Widodo, “Klasifikasi Penyakit Gigi Dan Mulut Menggunakan Metode Support Vector Klasifikasi Penyakit Gigi Dan Mulut Menggunakan Metode SVM,” no. January, 2018.
- [30] M. E. P. Ahmad Fikri Mauriza Yusuf Sulisty Nugroho, S.T., “IMPLEMENTASI DATA MINING UNTUK MEMPREDIKSI KELULUSAN MAHASISWA FAKULTAS KOMUNIKASI DAN INFORMATIKA UMS MENGGUNAKAN METODE NAIVE BAYES,” pp. 561–565, 2014.
- [31] P. Mata, “Analisis Perbandingan Metode Bayesian Network dan Dempster-

- Shafer Pada Sistem Pakar Diagnosis,” pp. 239–244, 2012.
- [32] R. W. Witjaksono, M. Wiyogo, and P. N. Wicaksono, “Perancangan Aplikasi Business Intelligence Pada Sistem Informasi Distribusi Pt Pertamina Lubricant Menggunakan Pentaho,” *J. Rekayasa Sist. Ind.*, vol. 2, no. 2, pp. 12–18, 2015.
- [33] H. A. Putranto, O. Setyawati, and A. L. Belakang, “Pengaruh Phrase Detection dengan POS -Tagger terhadap Akurasi Klasifikasi Sentimen menggunakan SVM,” vol. 5, no. 4, pp. 252–259, 2016.
- [34] A. Darmawan, N. Kustian, W. Rahayu, T. Tabebuya, and K. Pengunjung, “IMPLEMENTASI DATA MINING MENGGUNAKAN MODEL SVM,” vol. 2, no. 3, pp. 299–307, 2018.
- [35] S. Antar, B. Vol, V. I. No, and E. Supriyadi, “Metode SVM Berbasis PSO untuk Meningkatkan Prediksi Ketepatan Waktu Kelulusan Mahasiswa,” no. 2, pp. 113–120, 2017.
- [36] W. F. HARDIYANTI, “KLASIFIKASI DOKUMEN SITUS WEB BERITA ONLINE TENTANG KECELAKAAN LALU LINTAS MENGGUNAKAN NAÏVE BAYES CLASSIFIER (NBC) DAN SVM” 2019.
- [37] C.C.ASTUTI," Analisis Korelasi untuk Mengetahui Keeratan Hubungan antara Keaktifan Mahasiswa dengan Hasil Belajar Akhir "2017.
- [38] B. Santosa, “Tutorial Support Vector Machine 1 Ide Dasar Support Vector Machine,” pp. 1–23, 2011.

SURAT KEPUTUSAN DEKAN FTK UIN AR-RANIRY BANDA ACEH
NOMOR: B-2398.b/Un.08/FTK/KP.07.6/02/2020

TENTANG:

PENGANGKATAN PEMBIMBING SKRIPSI MAHASISWA FAKULTAS TARBİYAH DAN KEGURUAN
UIN AR-RANIRY BANDA ACEH

DEKAN FTK UIN AR-RANIRY BANDA ACEH

Menimbang : a. bahwa untuk kelancaran bimbingan skripsi dan ujian munaqasyah mahasiswa pada Fakultas Tarbiyah dan Keguruan UIN Ar-Raniry Banda Aceh maka dipandang perlu menunjuk pembimbing skripsi tersebut yang dituangkan dalam Surat Keputusan Dekan;
 b. bahwa saudara yang tersebut namanya dalam surat keputusan ini dipandang cakap dan memenuhi syarat untuk diangkat sebagai pembimbing skripsi.

Mengingat : 1. Undang-Undang Nomor 20 Tahun 2003, tentang Sistem Pendidikan Nasional;
 2. Undang-Undang Nomor 14 Tahun 2005, tentang Guru dan Dosen;
 3. Undang-Undang Nomor 12 Tahun 2012, tentang Sistem Pendidikan Tinggi;
 4. Peraturan Pemerintah No. 74 Tahun 2012 tentang Perubahan atas Peraturan Pemerintah RI Nomor 23 Tahun 2005 tentang Pengelolaan Keuangan Badan Layanan Umum;
 5. Peraturan Pemerintah Nomor 4 Tahun 2014 tentang Penyelenggaraan Pendidikan Tinggi dan Pengelolaan Perguruan Tinggi;
 6. Peraturan Presiden Nomor 64 Tahun 2013, tentang Perubahan Institut Agama Islam Negeri Ar-Raniry Banda Aceh menjadi Universitas Islam Negeri Ar-Raniry Banda Aceh;
 7. Peraturan Menteri Agama RI Nomor 12 Tahun 2014, tentang Organisasi & Tata Kerja UIN Ar-Raniry Banda Aceh;
 8. Peraturan Menteri Agama RI Nomor 21 Tahun 2015, tentang Statuta UIN Ar-Raniry Banda Aceh;
 9. Keputusan Menteri Agama Nomor 492 Tahun 2003, tentang Pendelegasian Wewenang Pengangkatan, Pemindahan, dan Pemberhentian PNS di Lingkungan Depag. RI;
 10. Keputusan Menteri Keuangan Nomor 293/KMK.05/2011 tentang Penetapan Institut Agama Islam Negeri Ar-Raniry Banda Aceh pada Kementerian Agama sebagai Instansi Pemerintah yang Menerapkan Pengelolaan Badan Layanan Umum;
 11. Keputusan Rektor UIN Ar-Raniry Nomor 01 Tahun 2015, tentang Pendelegasian Wewenang Kepada Dekan dan Direktur Pascasarjana di Lingkungan UIN Ar-Raniry Banda Aceh;

Memperhatikan : Keputusan Sidang/Seminar Proposal Skripsi Prodi Pendidikan Teknologi Informasi tanggal 13 Januari 2020

MEMUTUSKAN

Menetapkan
PERTAMA

Menunjuk Saudara:

1. Bustami, M.Sc sebagai pembimbing pertama
2. Hendri Ahmadian, S.Si., M.J.M sebagai pembimbing kedua

Untuk membimbing skripsi :

Nama : Marhamah
 NIM : 160212034
 Program Studi : Pendidikan Teknologi Informasi
 Judul Skripsi : Implementasi Algoritma Support Vector Machine (SVM) dalam Memprediksi Kelulusan Jalur Masuk Perguruan Tinggi Banda Aceh (Studi Kasus : Mahasiswa Baru Tahun Ajaran 2019)

KEDUA : Pembiayaan honorarium pembimbing pertama dan kedua tersebut di atas dibebankan pada DIPA UIN Ar-Raniry Banda Aceh Tahun 2020;

KETIGA : Surat Keputusan ini berlaku sampai akhir semester Ganjil Tahun Akademik 2020/2021

KEEMPAT : Surat Keputusan ini berlaku sejak tanggal ditetapkan dengan ketentuan bahwa segala sesuatu akan diubah dan diperbaiki kembali sebagaimana mestinya, apabila kemudian hari ternyata terdapat kekeliruan dalam surat keputusan ini.

Ditetapkan di : Banda Aceh
 Pada tanggal : 26 Februari 2020

An. Rektor
 Dekan


 Muslim Razali

Tembusan

1. Rektor UIN Ar-Raniry di Banda Aceh;
2. Ketua Prodi Pendidikan Teknologi Informasi;
3. Pembimbing yang bersangkutan untuk dimaklumi dan dilaksanakan;
4. Yang bersangkutan.

LAMPIRAN

Data

No	Variabel	No	Variabel
1	Jumlah saudara kandung	12	Prestasi di sekolah
2	Ortu yang dimiliki sekarang	13	Les privat di sekolah
3	Siapa yang membiayai sekolah	14	Berapa lama les privat
4	Pendidikan terakhir ayah	15	Bimbel seleksi masuk PTN
5	Pendidikan terakhir ibu	16	Berapa lama bimbel
6	Pekerjaan ayah	17	Nilai UN
7	Pekerjaan ibu	18	Jalur yang diikuti
8	Asal sekolah	19	Jalur minat
9	Jurusan di sekolah	20	Jalur lulus
10	Organisasi di sekolah	21	Sesuai atau tidak dengan jalur minat
11	Organisasi di luar sekolah	22	Jurusan lulus

U1	apakah sesuai atau tidak sesuai dengan jalur yang diminati																								
C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W					
Siapa yang Apa pend/ Apa pend/ Apa peker/ Apa peker/ kabupaten/ Apa jurusan/ Apakah ar/ Jika iya/ Apakah ar/ Apakah ar/ Apakah ar/ Sebelum/ r/ Jika iya/ Berapa nil/ Jalur mas/ Jalur yang/ Jalur apa/																				apakah se	Anda lulus di jur				
Bapak saja/ PERGURU/ PERGURU/ Pegawai	Tidak	0	Tidak	Tidak	Tidak	Tidak	0	99	SNMPTN/ SNMPTN/ UM-PTKIN	TIDAK SES PTI															
Ibu saja/ SMA/ SEDE/ PERGURU/ Sakit (Tid/ pegawai/ n/ Kota/ Band/ TKJ	Ya	2	Ya	Ya	Tidak	Tidak	0	46	SNMPTN/ SNMPTN/ PMB	TIDAK SES PTI															
Bapak & ibu/ PERGURU/ SMA/ SEDE/ Petani	Ibu rumah/ Bireuen	Teknik pe	Tidak	0	Tidak	Tidak	Tidak	0	76	SNMPTN/ SNMPTN/ PMB	TIDAK SES PTL														
Bapak saja/ PERGURU/ SMA/ SEDE/ Sopir	Ibu rumah/ Aceh Sela	IPA	Tidak	0	Ya	Ya	Ya	2	45	UM-PTKIN/ UM-PTKIN/ PMB	TIDAK SES PKM														
Bapak & ibu/ PERGURU/ PERGURU/ Petani	Petani	Bener me	Pa/	Tidak	0	Ya	Tidak	Tidak	0	52	SBMPTN/ FSNMPTN/ PMB	TIDAK SES PBL													
Abang dar/ SMA/ SEDE/ SMA/ SEDE/ Sudah me	pegawai/ n/ Banda/ Ace/ TKJ	Ya	2	Tidak	Ya	Ya	Tidak	0	50	SNMPTN/ SNMPTN/ SBMPTN	TIDAK SES PTL														
Ibu saja/ SMA/ SEDE/ SMP/ SEDE/ Pegawai/ n/ pegawai/ n/ Aceh/ teng/ IPA	Tidak	0	Ya	Ya	Tidak	Tidak	0	41	UM-PTKIN/ UM-PTKIN/ PMB	TIDAK SES TEKNIK PERTANI															
Bapak & ibu/ PERGURU/ PERGURU/ Wiraswas/ Ibu rumah/ Aceh/ bes/ IPA	Tidak	0	Tidak	Ya	Tidak	Tidak	0	40	SBMPTN/ SBMPTN/ SBMPTN	SESUAI/ TEKNIK LINGKUN															
Ibu saja/ SMA/ SEDE/ PERGURU/ BUMN	Ibu rumah/ Pidie	IPA	Tidak	0	Ya	Ya	Tidak	0	85	SNMPTN/ SNMPTN/ UM-PTKIN	TIDAK SES HUKUM EKONOMI														
Bapak & ibu/ SMA/ SEDE/ SMA/ SEDE/ Petani	Ibu rumah/ Pidie/ Jaya/ IPA	Tidak	2	Ya	Tidak	Tidak	Tidak	2	48	SNMPTN/ SNMPTN/ PMB	TIDAK SES PTL														
Bapak saja/ SMA/ SEDE/ PERGURU/ Pegawai/ n/ Ibu rumah/ Aceh/ Bara/ IPA	Tidak	0	Tidak	Ya	Ya	Tidak	0	65	SNMPTN/ SNMPTN/ PMB	TIDAK SES LTE															
Bapak saja/ PERGURU/ PERGURU/ Wiraswas/ Ibu rumah/ Aceh/ teng/ TKJ	Tidak	0	Tidak	Ya	Tidak	Ya	1	87	SNMPTN/ SNMPTN/ SBMPTN	TIDAK SES ILMU PERUSTAH															
Ibu saja/ SMA/ SEDE/ SMA/ SEDE/ Wiraswas/ pegawai/ n/ Aceh/ selat/ IPA	Tidak	0	Ya	Ya	Ya	Tidak	0	56	SBMPTN/ SBMPTN/ UM-PTKIN	TIDAK SES MPI															
Bapak & ibu/ TIDAK/ SEK/ SMA/ SEDE/ Pegawai/ n/ pegawai/ n/ SMA/ MU/ IPA	Tidak	0	Ya	Tidak	Tidak	Tidak	0	33	SNMPTN/ SNMPTN/ SNMPTN	SESUAI/ MPI															
Bapak & ibu/ SMA/ SEDE/ SMA/ SEDE/ Pegawai/ n/ pegawai/ n/ Aceh/ bes/ IPA	Tidak	0	Ya	Ya	Tidak	Ya	Tidak	0	55	SBMPTN/ SBMPTN/ SBMPTN	SESUAI/ TEKNIK LINGKUN														
Bapak & ibu/ PERGURU/ SMA/ SEDE/ Wiraswas/ pegawai/ n/ banda/ ace/ IPA	Ya	2	Ya	Ya	Ya	Tidak	2	76	PMB/ PMB/ PMB	SESUAI/ ILMU EKONOMI															
Bapak & ibu/ PERGURU/ SMA/ SEDE/ Petani	Ibu rumah/ Aceh/ Teng/ IPS	Tidak	0	Tidak	Ya	Tidak	Tidak	0	50	SPAN-PTK/ SPAN-PTK/ PMB	TIDAK SES PKF														
Bapak saja/ TIDAK/ SEK/ SMP/ SEDE/ Wiraswas/ Ibu rumah/ Bireuen/ IPA	Tidak	0	Ya	Ya	Ya	Tidak	0	55	SBMPTN/ SBMPTN/ PMB	TIDAK SES PAI															
Bapak saja/ SMA/ SEDE/ SMA/ SEDE/ Tani	Ibu rumah/ Bireuen/ IPA	Ya	2	Tidak	Tidak	Tidak	Tidak	0	90	SPAN-PTK/ SPAN-PTK/ PMB	TIDAK SES TEKNIK LINGKUN														
Bapak saja/ SMA/ SEDE/ PERGURU/ Wiraswas/ Ibu rumah/ Aceh/ bes/ IPA	Tidak	0	Tidak	Ya	Tidak	Tidak	0	51	SBMPTN/ SBMPTN/ PMB	TIDAK SES TEKNIK LINGKUN															
Bapak saja/ SD/ SEDER/ SD/ SEDER/ Tni al/ Ibu rumah/ Kota/ surat/ IPS	Tidak	0	Tidak	Tidak	Tidak	Tidak	0	54	SNMPTN/ SNMPTN/ SBMPTN	TIDAK SES PTI															
Bapak saja/ SMA/ SEDE/ SD/ SEDER/ Buruh/ Kas/ Ibu rumah/ Aceh/ IPS	Tidak	0	Tidak	Ya	Tidak	Ya	4	87	SNMPTN/ SNMPTN/ SNMPTN	SESUAI/ PKF															
Orang lain/ SMA/ SEDE/ PERGURU/ Sudah mn/ Wiraswas/ Banda/ ace/ IPA	Tidak	0	Tidak	Tidak	Ya	Tidak	0	60	UM-PTKIN/ UM-PTKIN/ UM-PTKIN	SESUAI/ Bimbingan dan k															
Bapak & ibu/ SMP/ SEDE/ SMP/ SEDE/ Pedagang/ Ibu rumah/ Pidie/ IPA	Tidak	0	Tidak	Tidak	Tidak	Tidak	0	71	SPAN-PTK/ SPAN-PTK/ UM-PTKIN	TIDAK SES Administrasi neg															
Bapak & ibu/ SMP/ SEDE/ SEDER/ Petani	Mengurus/ Simeulue/ IPA	Tidak	0	Tidak	Ya	Ya	Tidak	0	87	SNMPTN/ SNMPTN/ PMB	TIDAK SES Perbankan syaria														
Bapak & ibu/ PERGURU/ PERGURU/ Petani	Ibu rumah/ Aceh/ Huta/ IPA	Tidak	0	Tidak	Tidak	Tidak	0	60	NM/ NM/ NM	CECILIA/ DVC															

```

File Edit Format View Help
Berapa jumlah saudara kandung anda?,orang tua yang anda miliki sekarang?,siapa yang membayai sekolah anda?,apa pendidikan terakhir ayah anda?,apa pendidikan terakhir i
1,ayah & ibu,bapak saja,PERGURUAN TINGGI,PERGURUAN TINGGI,Pegawai BUMN,Ibu rumah tangga,kota Langsa,IPA,Tidak,0,Tidak,Tidak,Tidak,0,99,"SNMPTN,SBMPTN, SPAN-PTKIN,
2,ayah & ibu,Ibu saja,SMA/SEDERAJAT,PERGURUAN TINGGI,Sakit (Tidak Bekerja),pegawai negeri,kota Banda Aceh,TKJ,Ya,2,Ya,Ya,Tidak,Tidak,0,46,"SNMPTN,SBMPTN, UM-PTKIN, PMB"
1,ayah & ibu,bapak & ibu,PERGURUAN TINGGI,SMA/SEDERAJAT,Petani,Ibu rumah tangga,Bireuen,Teknik pemesian,Tidak,0,Tidak,Tidak,Tidak,0,76,"SNMPTN,SBMPTN, UM-PTKIN,
4,ayah & ibu,bapak saja,PERGURUAN TINGGI,SMA/SEDERAJAT,Sopir,Ibu rumah tangga,Aceh Selatan,IPA,Tidak,0,Ya,Ya,Ya,2,45,"UM-PTKIN, PMB",UM-PTKIN, PMB,TIDAK SESUAI,PMB
4,ayah & ibu,bapak & ibu,PERGURUAN TINGGI,PERGURUAN TINGGI,Petani,Petani,Bener meriah,Pa/,Tidak,0,Ya,Tidak,Tidak,Tidak,0,52,"SBMPTN, PMB",SBMPTN, PMB,TIDAK SESUAI,PBL
1,Ibu saja (ayah sudah meninggal),Abang dan ibu,SMA/SEDERAJAT,SMA/SEDERAJAT,Sudah meninggal,pegawai negeri,banda Aceh,TKJ,Ya,2,Tidak,Ya,Ya,Tidak,0,50,"SNMPTN,SBMPTN",S
1,ayah & ibu,Ibu saja,SMA/SEDERAJAT,SMP/SEDERAJAT,Pegawai negeri,pegawai negeri,Aceh Tenggara,IPA,Tidak,0,Ya,Ya,Tidak,Tidak,0,41,"UM-PTKIN, PMB",UM-PTKIN, PMB,TIDAK SESU
3,ayah & ibu,bapak & ibu,PERGURUAN TINGGI,PERGURUAN TINGGI,Wiraswasta,Ibu rumah tangga,Aceh Besar,IPA,Tidak,0,Tidak,Ya,Tidak,Tidak,0,65,"SNMPTN,SPAN-PTKIN,SBMPTN,UM-P
1,ayah & ibu,bapak & ibu,PERGURUAN TINGGI,PERGURUAN TINGGI,Petani,Petani,Bener meriah,Pa/,Tidak,0,Ya,Tidak,Tidak,Tidak,0,52,"SBMPTN, PMB",SBMPTN, PMB,TIDAK SESUAI,PBL
2,Ibu saja (ayah sudah meninggal),Ibu saja,SMA/SEDERAJAT,PERGURUAN TINGGI,BUMN,Ibu rumah tangga,Pidie,IPA,Tidak,0,Ya,Ya,Tidak,Tidak,0,85,"SNMPTN,SBMPTN, SBMPTN,SPAN-PT
1,ayah & ibu,bapak & ibu,SMA/SEDERAJAT,SMA/SEDERAJAT,Petani,Ibu rumah tangga,Pidie Jaya,IPA,Tidak,2,Ya,Tidak,Tidak,Tidak,2,48,"SNMPTN,SBMPTN, SBMPTN,SPAN-PTKIN,UM-PTKI
2,ayah & ibu,bapak & ibu,SMA/SEDERAJAT,PERGURUAN TINGGI,Pegawai negeri,Ibu rumah tangga,Aceh Barat Daya,IPA,Tidak,0,Tidak,Ya,Tidak,Tidak,0,65,"SNMPTN,SPAN-PTKIN,SBMPTN,UM-P
1,ayah & ibu,bapak & ibu,SMA/SEDERAJAT,SMA/SEDERAJAT,Pegawai negeri,Ibu rumah tangga,Aceh Tengah,TKJ,Tidak,0,Tidak,Ya,Tidak,Ya,1,87,"SNMPTN,SBMPTN, SBMPTN,SPAN-PTKIN",
2,ayah & ibu,Ibu saja,SMA/SEDERAJAT,SMA/SEDERAJAT,Wiraswasta,pegawai negeri,Aceh Selatan,IPA,Tidak,0,Ya,Ya,Ya,Tidak,0,56,"SBMPTN,SBMPTN, SBMPTN,SPAN-PTKIN,UM-PTKI
3,ayah & ibu,bapak & ibu,SMA/SEDERAJAT,SMA/SEDERAJAT,Petani,Ibu rumah tangga,Aceh Tenggara,IPS,Tidak,0,Tidak,Ya,Tidak,Tidak,0,33,"SBMPTN,SBMPTN, SBMPTN,UM-PTKI
1,ayah & ibu,bapak & ibu,PERGURUAN TINGGI,SMA/SEDERAJAT,Petani,Ibu rumah tangga,Aceh Tenggara,IPS,Tidak,0,Tidak,Ya,Tidak,Tidak,0,33,"SBMPTN,SBMPTN, SBMPTN,UM-PTKI
3,ayah & ibu,bapak & ibu,TIDAK SEKOLAH,SMP/SEDERAJAT,Wiraswasta,Ibu rumah tangga,Bireuen,IPA,Tidak,0,Ya,Ya,Ya,Tidak,0,55,"SBMPTN, PMB",SBMPTN, PMB,TIDAK SESUAI,PAI
8,ayah & ibu,bapak & ibu,SMA/SEDERAJAT,SMA/SEDERAJAT,Tani,Ibu rumah tangga,Bireuen,IPA,Ya,2,Tidak,Tidak,Tidak,Tidak,0,90,"SPAN-PTKIN, UM-PTKIN, PMB",SBMPTN, PMB,TIDAK SESU
1,ayah & ibu,bapak & ibu,SMA/SEDERAJAT,PERGURUAN TINGGI,Wiraswasta,Ibu rumah tangga,Aceh Besar,IPA,Tidak,0,Tidak,Ya,Tidak,Tidak,0,51,"SBMPTN, PMB",SBMPTN, PMB,TIDAK SESU
2,ayah & ibu,bapak & ibu,SD/SEDERAJAT,SD/SEDERAJAT,Tni al,Ibu rumah tangga,kota surabaya,IPS,Tidak,0,Tidak,Tidak,Tidak,0,34,"SNMPTN,SBMPTN, SBMPTN,SBMPTN,TIDAK SES
1,ayah & ibu,bapak & ibu,SMA/SEDERAJAT,SD/SEDERAJAT,Buruh Kasar,Ibu rumah tangga,Aceh,IPS,Tidak,0,Tidak,Ya,Tidak,Ya,4,87,"SNMPTN,SNMPTN,SNMPTN,SESUAI,PKF
1,Ibu saja (ayah sudah meninggal),Orang lain/wali,SMA/SEDERAJAT,PERGURUAN TINGGI,Sudah meninggal,Wiraswasta,banda Aceh,IPA,Tidak,0,Tidak,Tidak,Ya,Tidak,0,60,"UM-PTKIN,
3,ayah & ibu,bapak & ibu,SMP/SEDERAJAT,SMP/SEDERAJAT,Pedagog Diasa,Ibu rumah tangga,Pidie,IPA,Tidak,0,Tidak,Tidak,Tidak,0,71,"SPAN-PTKIN, SBMPTN,SBMPTN,SPAN-PTKIN,UM
4,ayah & ibu,bapak & ibu,SMP/SEDERAJAT,SD/SEDERAJAT,Petani,Mengurus rumah tangga,Simeulue Barat,IPA,Tidak,0,Tidak,Ya,Ya,Tidak,0,87,"SNMPTN,SBMPTN, SBMPTN,PMB,TIDAK SESU
2,ayah & ibu,bapak & ibu,PERGURUAN TINGGI,PERGURUAN TINGGI,Petani,Ibu rumah tangga,Aceh Timur,IPA,Tidak,0,Tidak,Tidak,Tidak,0,58,"PMB, PMB, PMB,SESUAI,PKF
3,ayah & ibu,bapak & ibu,SMP/SEDERAJAT,SMP/SEDERAJAT,Wiraswasta,Wiraswasta,kab.Bener Meriah,TKJ,Ya,2,Tidak,Ya,Ya,2,54,"SBMPTN,SBMPTN, SBMPTN,SPAN-PTKIN,UM-PTKI
1,ayah & ibu,bapak & ibu,SD/SEDERAJAT,SD/SEDERAJAT,Poliisi,Ibu rumah tangga,banda Aceh,IPS,Ya,2,Tidak,Ya,Ya,2,62,"SNMPTN,SBMPTN, SBMPTN,SBMPTN,TIDAK SESUAI,Hukum Ekon
3,ayah & ibu,bapak & ibu,PERGURUAN TINGGI,SMA/SEDERAJAT,Petani,Ibu rumah tangga,banda Aceh,IPS,Tidak,0,Tidak,Ya,Tidak,Tidak,0,48,"UM-PTKIN, PMB",UM-PTKIN, PMB,TIDAK SESU
5,ayah & ibu,bapak & ibu,PERGURUAN TINGGI,PERGURUAN TINGGI,Buruh,buruh,banda Aceh,IPS,Ya,3,Ya,Tidak,Tidak,Ya,3,87,"SNMPTN,SBMPTN, SBMPTN,SESUAI,Teknik Lingkungan
4,ayah & ibu,Ibu saja,SMA/SEDERAJAT,SMA/SEDERAJAT,Wiraswasta,pegawai negeri,Kabupaten Bireuen,Teknik otomotif TKR,Tidak,0,Tidak,Ya,Ya,Tidak,0,29,"UM-PTKIN, PMB",UM-PTKI
3,ayah & ibu,bapak & ibu,SMP/SEDERAJAT,SMP/SEDERAJAT,Pegawai negeri,pegawai negeri,Aceh Besar,IPS,Ya,2,Ya,Tidak,Tidak,Ya,2,60,"SBMPTN,SPAN-PTKIN, UM-PTKIN, PMB",SBMPTN,
4,ayah & ibu,bapak & ibu,SMA/SEDERAJAT,SMA/SEDERAJAT,Pegawai negeri,Ibu rumah tangga,Aceh,IPS,Ya,1,Tidak,Ya,Tidak,Tidak,1,35,"SNMPTN,SBMPTN, SBMPTN,SESUAI,H
2,Ibu saja (ayah sudah meninggal),Ibu saja,PERGURUAN TINGGI,PERGURUAN TINGGI,Buruh,pegawai negeri,Kabupaten solok,IPA,Tidak,0,Ya,Ya,Tidak,0,66,"SBMPTN,SPAN-PTKIN,UM
0,ayah & ibu,bapak & ibu,PERGURUAN TINGGI,TIDAK SEKOLAH,Buruh,Ibu rumah tangga,Kabupaten Aceh Barat Daya,IPA,Tidak,0,Tidak,Ya,Tidak,Tidak,0,55,"SBMPTN,SPAN-PTKIN,UM-PT
1,ayah & ibu,bapak & ibu,PERGURUAN TINGGI,PERGURUAN TINGGI,Wiraswasta,Ibu rumah tangga,Aceh Besar,IPA,Tidak,0,Tidak,Ya,Tidak,Ya,3,60,"SPAN-PTKIN, UM-PTKIN, PMB",UM-PTKIN, PMB
1,Ibu & ayah sudah meninggal,Orang lain/wali,SMA/SEDERAJAT,SMP/SEDERAJAT,Tidak ada pekerjaan tetap,Ibu rumah tangga,banda Aceh,IPA,Tidak,0,Ya,Ya,Tidak,0,53,"UM-PT
1,Ibu saja (ayah sudah meninggal),Ibu saja,SMP/SEDERAJAT,SMA/SEDERAJAT,Pegawai negeri,pegawai negeri,Aceh Besar,KEAGAMAAN,Tidak,0,Tidak,Ya,Tidak,0,78,"UMPTKIN,UM-PT
1,ayah & ibu,Ibu saja,PERGURUAN TINGGI,PERGURUAN TINGGI,Buruh,pegawai negeri,Aceh Selatan,IPA,Tidak,0,Ya,Ya,Ya,3,60,"SPAN-PTKIN, UM-PTKIN, PMB",UM-PTKIN, PMB,TIDAK SESU
2,ayah & ibu,bapak & ibu,PERGURUAN TINGGI,PERGURUAN TINGGI,Pegawai negeri,pegawai negeri,Bireuen,IPA,Tidak,0,Tidak,Tidak,Tidak,0,55,"SNMPTN,SBMPTN, SBMPTN,SPAN-PTKIN,UM
0,ayah & ibu,bapak & ibu,SMP/SEDERAJAT,PERGURUAN TINGGI,Pegawai negeri,pegawai negeri,Aceh Selatan,IPA,Tidak,0,Ya,Ya,Tidak,0,83,"SNMPTN,SPAN-PTKIN, UM-PTKIN, PMB",
1,ayah & ibu,bapak & ibu,SMP/SEDERAJAT,SMP/SEDERAJAT,Petani,Ibu rumah tangga,Pidie,IPS,Ya,4,Ya,Ya,Ya,Tidak,0,49,"PMB, PMB, PMB,SESUAI,PAI
2,ayah & ibu,bapak & ibu,PERGURUAN TINGGI,PERGURUAN TINGGI,Pegawai negeri,Wiraswasta,Aceh Tenggara,IPS,Tidak,2,Tidak,Tidak,Tidak,Tidak,2,63,"SPAN-PTKIN, UM-PTKIN",SPAN-
3,ayah & ibu,bapak & ibu,PERGURUAN TINGGI,PERGURUAN TINGGI,Pegawai negeri,Tidak ada,Aceh Besar,IPA,Tidak,0,Ya,Tidak,Tidak,0,42,"SNMPTN,SBMPTN, SBMPTN,SPAN-PTKIN, SPAN
1,Ibu saja (ayah sudah meninggal),Ibu saja,SMA/SEDERAJAT,PERGURUAN TINGGI,Wiraswasta,Wiraswasta,Kabupaten Aceh Besar,IPA,Tidak,0,Tidak,Ya,Tidak,Tidak,2,60,"SBMPTN,SPAN
2,ayah & ibu,bapak & ibu,SMA/SEDERAJAT,SMA/SEDERAJAT,Petani,Aceh Tengah,IPS,Tidak,2,Ya,Ya,Ya,Tidak,0,60,"UM-PTKIN, PMB",UM-PTKIN, PMB,TIDAK SESUAI, SOSIOLOGI agana
5,ayah & ibu,Ibu saja,SMA/SEDERAJAT,SMA/SEDERAJAT,Petani,Aceh Tenggara,Persekerjaan,TKJ,Tidak,0,Ya,Ya,Ya,2,53,"UM-PTKIN, PMB",UM-PTKIN, PMB,TIDAK SESUAI,KESEJA
2,ayah & ibu,Ibu sudah meninggal),bapak & ibu,SMA/SEDERAJAT,SMP/SEDERAJAT,Pegawai negeri,Ibu rumah tangga,Pijay,IPA,Tidak,0,Ya,Tidak,Ya,Ya,1,77,"SPAN-PTKIN, UM-PTKIN, U

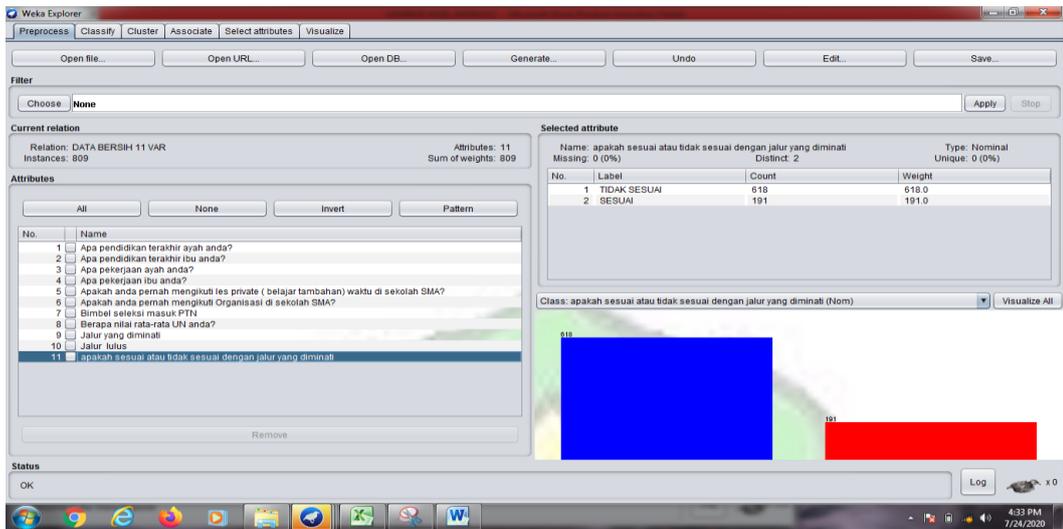
```

```

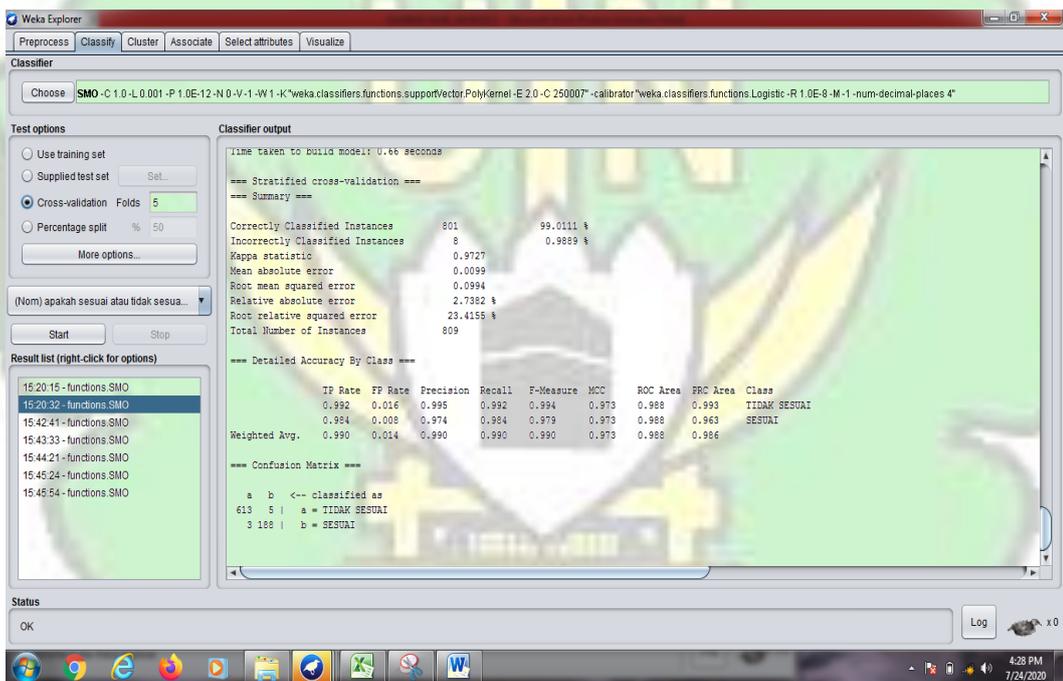
Relation 'DATA BARU 810'
@attribute 'Berapa Jumlah saudara kandung anda?' numeric
@attribute 'Orang tua yang anda miliki sekarang?' {'Ayah & ibu','Ibu saja (ayah sudah meninggal)','ibu & ayah sudah meninggal','Ayah saja (ibu sudah meninggal)','Ayah dan ibu','Ayah & ibu','Ayah dan ibu tetapi sudah cerai','Cerai','Tinggal sama ayah','Ayah & ibu, Ibu dan ayah bercerai',Ibu}
@attribute 'Siapa yang membayai sekolah anda?' {'Bapak saja','Ibu saja','Bapak & ibu','Abang dan ibu','Orang lain/wali','Bapak & ibu, Ibu saja','biaya sendiri','Bapak & ibu, biaya sendiri','nenek','Abang Ipar','Bapak & ibu, Orang lain/wali','Nenek,Kakak}
@attribute 'Apa pendidikan terakhir ayah anda?' {'PERGURUAN TINGGI',SMA/SEDERAJAT,'TIDAK SEKOLAH',SD/SEDERAJAT,SMP/SEDERAJAT}
@attribute 'Apa pendidikan terakhir ibu anda?' {'PERGURUAN TINGGI',SMA/SEDERAJAT,SMP/SEDERAJAT,SD/SEDERAJAT,'TIDAK SEKOLAH'}
@attribute 'Apa pekerjaan ayah anda?' {'Pegawai BUMN','Sakit (Tidak Bekerja)','Petani,Sopir','Sudah meninggal','Pegawai negeri','Wiraswasta,BUMN,Tani','Tni al','Buruh Kasar','Sudah

```

Praproses Pada Weka



Pengujian Dengan Cross-Validatio Dan Percentage Split



Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier: Choose SMO-C 1.0-L 0.001-P 1.0E-12-N 0-V-1-W 1-K/weka.classifiers.functions.supportVector.PolyKernel-E 2.0-C 25000*-calibrator/weka.classifiers.functions.Logistic-R 1.0E-8-M 1-num-decimal-places 4

Test options:

- Use training set
- Supplied test set
- Cross-validation Folds 10
- Percentage split % 50

Classifier output:


```
Time taken to build model: 0.7 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances      802      99.1347 %
Incorrectly Classified Instances     7       0.8653 %
Kappa statistic                    0.9761
Mean absolute error                 0.0087
Root mean squared error             0.093
Relative absolute error             2.3962 %
Root relative squared error        21.9032 %
Total Number of Instances          809

=== Detailed Accuracy By Class ===
          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
Weighted Avg.   0.991  0.010  0.991    0.991  0.991    0.976  0.991  0.988  SESUAI

=== Confusion Matrix ===
  a  b  <-- classified as
613  5  | a = TIDAK SESUAI
 2 159 | b = SESUAI
```

Status: OK

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier: Choose SMO-C 1.0-L 0.001-P 1.0E-12-N 0-V-1-W 1-K/weka.classifiers.functions.supportVector.PolyKernel-E 2.0-C 25000*-calibrator/weka.classifiers.functions.Logistic-R 1.0E-8-M 1-num-decimal-places 4

Test options:

- Use training set
- Supplied test set
- Cross-validation Folds 10
- Percentage split % 20

Classifier output:


```
Time taken to test model on test split: 0.02 seconds
=== Summary ===
Correctly Classified Instances      616      95.2087 %
Incorrectly Classified Instances     31       4.7913 %
Kappa statistic                    0.8652
Mean absolute error                 0.0479
Root mean squared error             0.2189
Relative absolute error             13.1546 %
Root relative squared error        51.6192 %
Total Number of Instances          647

=== Detailed Accuracy By Class ===
          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
Weighted Avg.   0.952  0.097  0.952    0.952  0.952    0.865  0.928  0.929  SESUAI

=== Confusion Matrix ===
  a  b  <-- classified as
482  13 | a = TIDAK SESUAI
 18 134 | b = SESUAI
```

Status: OK

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose SMO-C 1.0-L 0.001-P 1.0E-12-N 0-V-1-W 1-K-weka.classifiers.functions.supportVector.PolyKernel-E 2.0-C 250007-calibrator-weka.classifiers.functions.Logistic-R 1.0E-8-M-1-num-decimal-places 4

Test options

Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 30
 More options...

(Nom) apakah sesuai atau tidak sesuai...

Start Stop

Result list (right-click for options)

- 15:20:15 - functions.SMO
- 15:20:32 - functions.SMO
- 15:42:41 - functions.SMO
- 15:43:33 - functions.SMO
- 15:44:21 - functions.SMO
- 15:45:24 - functions.SMO
- 15:45:54 - functions.SMO

Classifier output

Time taken to test model on test split: 0.02 seconds

=== Summary ===

Correctly Classified Instances	557	98.4099 %
Incorrectly Classified Instances	9	1.5901 %
Kappa statistic	0.9561	
Mean absolute error	0.0159	
Root mean squared error	0.1261	
Relative absolute error	4.3698 %	
Root relative squared error	29.8079 %	
Total Number of Instances	566	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
Weighted Avg.	0.984	0.015	0.995	0.984	0.990	0.956	0.984	0.992	TIDAK SESUAI
	0.985	0.016	0.949	0.985	0.967	0.956	0.984	0.938	SESUAI
Weighted Avg.	0.984	0.015	0.985	0.984	0.984	0.956	0.984	0.979	

=== Confusion Matrix ===

```

a b <-- classified as
427 7 | a = TIDAK SESUAI
2 100 | b = SESUAI
  
```

Status

OK Log x 0

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose SMO-C 1.0-L 0.001-P 1.0E-12-N 0-V-1-W 1-K-weka.classifiers.functions.supportVector.PolyKernel-E 2.0-C 250007-calibrator-weka.classifiers.functions.Logistic-R 1.0E-8-M-1-num-decimal-places 4

Test options

Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 40
 More options...

(Nom) apakah sesuai atau tidak sesuai...

Start Stop

Result list (right-click for options)

- 15:20:15 - functions.SMO
- 15:20:32 - functions.SMO
- 15:42:41 - functions.SMO
- 15:43:33 - functions.SMO
- 15:44:21 - functions.SMO
- 15:45:24 - functions.SMO
- 15:45:54 - functions.SMO

Classifier output

Time taken to test model on test split: 0.02 seconds

=== Summary ===

Correctly Classified Instances	461	99.1753 %
Incorrectly Classified Instances	4	0.8247 %
Kappa statistic	0.9772	
Mean absolute error	0.0082	
Root mean squared error	0.0908	
Relative absolute error	2.2758 %	
Root relative squared error	21.4775 %	
Total Number of Instances	465	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
Weighted Avg.	0.989	0.009	1.000	0.989	0.995	0.977	0.995	0.997	TIDAK SESUAI
	1.000	0.011	0.966	1.000	0.983	0.977	0.985	0.966	SESUAI
Weighted Avg.	0.992	0.009	0.992	0.992	0.992	0.977	0.995	0.990	

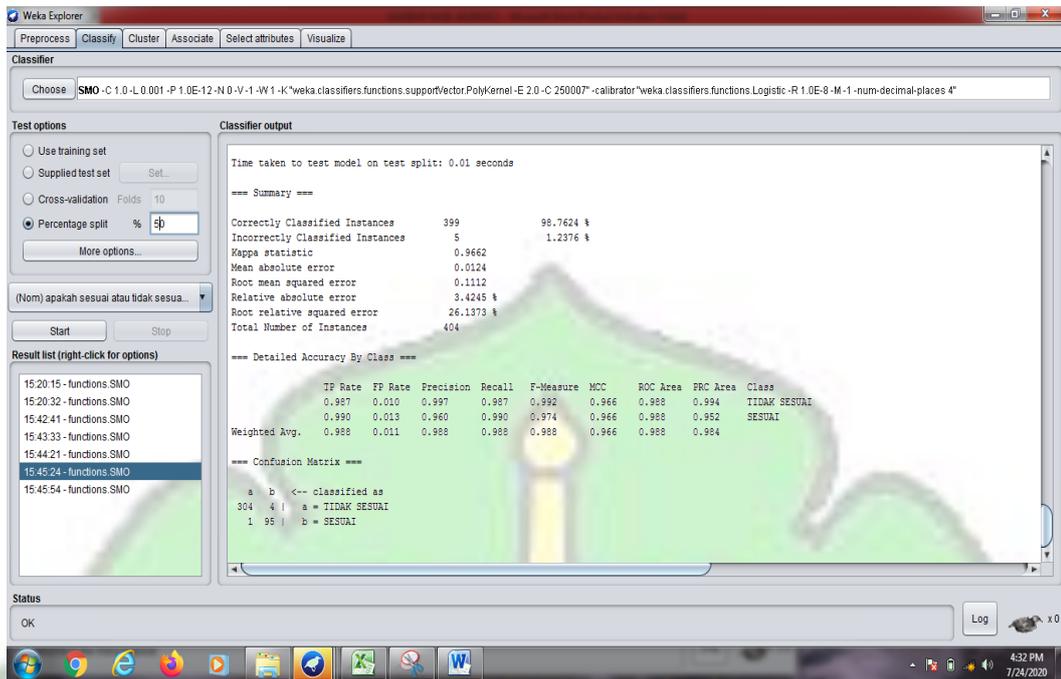
=== Confusion Matrix ===

```

a b <-- classified as
369 4 | a = TIDAK SESUAI
0 113 | b = SESUAI
  
```

Status

OK Log x 0



Akurasi dan Confussion Matrix

Time taken to build model: 1.14 seconds

=== Stratified cross-validation ===
 === Summary ===

```

Correctly Classified Instances      802      99.1347 %
Incorrectly Classified Instances     7         0.8653 %
Kappa statistic                    0.9761
Mean absolute error                 0.0087
Root mean squared error             0.093
Relative absolute error             2.3962 %
Root relative squared error        21.9032 %
Total Number of Instances          809

```

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
TIDAK SESUAI	0.992	0.010	0.997	0.992	0.994	0.976	0.991	0.995	TIDAK SESUAI
SESUAI	0.990	0.008	0.974	0.990	0.982	0.976	0.991	0.966	SESUAI
Weighted Avg.	0.991	0.010	0.991	0.991	0.991	0.976	0.991	0.988	

=== Confusion Matrix ===

```

  a  b  <-- classified as
613  5 | a = TIDAK SESUAI
 2 189 | b = SESUAI

```


Korelasi jalur dengan jalur lulus

		Jalur minat	Jalur lulus
Jalur minat	Pearson Correlation	1	.311**
	Sig. (2-tailed)		.000
	N	809	809
Jalur lulus	Pearson Correlation	.311**	1
	Sig. (2-tailed)	.000	
	N	809	809

** . Correlation is significant at the 0.01 level (2-tailed).

Korelasi bimbel seleksi masuk PTN dengan jalur lulus

		Bimbel seleksi masuk PTN	Jalur lulus
Bimbel seleksi masuk PTN	Pearson Correlation	1	-.108**
	Sig. (2-tailed)		.002
	N	809	809
Jalur lulus	Pearson Correlation	-.108**	1
	Sig. (2-tailed)	.002	
	N	809	809

** . Correlation is significant at the 0.01 level (2-tailed).

Korelasi nilai UN dengan jalur lulus

		Nilai UN	Jalur lulus
Nilai UN	Pearson Correlation	1	.014
	Sig. (2-tailed)		.684
	N	809	809
Jalur lulus	Pearson Correlation	.014	1
	Sig. (2-tailed)	.684	
	N	809	809

Korelasi organisasi di SMA dengan jalur lulus

		Organisasi di SMA	Jalur lulus
Organisasi di SMA	Pearson Correlation	1	.036
	Sig. (2-tailed)		.304
	N	809	809
Jalur lulus	Pearson Correlation	.036	1
	Sig. (2-tailed)	.304	
	N	809	809

Korelasi les private di SMA dengan jalur lulus

		Les private di SMA	Jalur lulus
Les private di SMA	Pearson Correlation	1	-.018
	Sig. (2-tailed)		.607
	N	809	809
Jalur lulus	Pearson Correlation	-.018	1
	Sig. (2-tailed)	.607	
	N	809	809

Korelasi pekerjaan ayah dengan jalur lulus

		Pekerjaan ayah	Jalur lulus
Pekerjaan ayah	Pearson Correlation	1	.086
	Sig. (2-tailed)		.015
	N	809	809
Jalur lulus	Pearson Correlation	.086	1
	Sig. (2-tailed)	.015	
	N	809	809

*. Correlation is significant at the 0.05 level (2-tailed).

Korelasi pekerjaan ibu dengan jalur lulus

		Pekerjaan ibu	Jalur lulus
Pekerjaan ibu	Pearson Correlation	1	.045
	Sig. (2-tailed)		.203
	N	809	809
Jalur lulus	Pearson Correlation	.045	1
	Sig. (2-tailed)	.203	
	N	809	809

Korelasi pendidikan terakhir ayah dengan jalur lulus

		Pendidikan terakhir ayah	Jalur lulus
Pendidikan terakhir ayah	Pearson Correlation	1	.024
	Sig. (2-tailed)		.497
	N	809	809
Jalur lulus	Pearson Correlation	.024	1
	Sig. (2-tailed)	.497	
	N	809	809

Korelasi pendidikan terakhir ibu dengan jalur lulus

		Pendidikan terakhir ibu	Jalur lulus
Pendidikan terakhir ibu	Pearson Correlation	1	-.015
	Sig. (2-tailed)		.667
	N	809	809
Jalur lulus	Pearson Correlation	-.015	1
	Sig. (2-tailed)	.667	
	N	809	809

Korelasi orang tua yang dimiliki dengan jalur lulus

		Orang tua yang dimiliki sekarang	Jalur lulus
Orang tua yang dimiliki sekarang	Pearson Correlation	1	.049
	Sig. (2-tailed)		.165
	N	809	809
Jalur lulus	Pearson Correlation	.049	1
	Sig. (2-tailed)	.165	
	N	809	809

Pengujian akurasi

Tabel 4. 11 Perbanding uji variabel

Pengujian	Variabel			
	Variabel Asli	Variabel Sesudah Cleaning	Variabel Korelasi	Tanpa Variabel korelasi
SVM	98.51 %	99.13 %	99.75%	74.53%

Cross Validation

Cross Validation	Akurasi	SVM
5	Precision	97%
	Recall	98%
	F-Measure	98%
10	Precision	97%
	Recall	99%
	F-Measure	98%

Percentage Split

Percentage Split	Akurasi	SVM
20%	Precision	95%
	Recall	95%
	F-Measure	95%
30%	Precision	99%
	Recall	98%
	F-Measure	98%
40%	Precision	99%
	Recall	99%
	F-Measure	99%
50%	Precision	99%
	Recall	99%
	F-Measure	99%

Perbandingan Jenis Pengujian

