ANALISIS KLASIFIKASI WAKTU TUNGGU KERJA DENGAN METODE *SUPPORT VECTOR MACHINE* DAN *NAÏVE BAYES CLASSIFICATION*

(Studi Kasus : Waktu tunggu kerja Alumni Universitas Islam Indonesia)

TUGAS AKHIR



Disusun Oleh:

Gustiara Dayu Amrinda

14 611 163

PROGRAM STUDI STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS ISLAM INDONESIA

YOGYAKARTA

2018

ANALISIS KLASIFIKASI WAKTU TUNGGU KERJA DENGAN METODE *SUPPORT VECTOR MACHINE* DAN *NAÏVE BAYES CLASSIFICATION*

(Studi Kasus : Waktu tunggu kerja Alumni Universitas Islam Indonesia)

TUGAS AKHIR

(Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Program Studi Statistika)



Gustiara Dayu Amrinda

14 611 163

PROGRAM STUDI STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS ISLAM INDONESIA

YOGYAKARTA

2018

HALAMAN PERSETUJUAN PEMBIMBING

TUGAS AKHIR

Judul : Analisis Klasifikasi Waktu Tunggu Kerja dengan Metode *Support Vector Machine* dan *Naïve Bayes Classification* (Studi Kasus : Waktu tunggu kerja Alumni Universitas Islam Indonesia)

Nama : Gustiara Dayu Amrinda

Nomor Mahasiswa : 14 611 163

TUGAS AKHIR INI TELAH DIPERIKSA DAN DISETUJUI UNTUK DIUJIKAN

Yogyakarta, 10 April 2018

Mengetahui,
Dosen Pembimbing

Ayundyah Kesumawati, S.Si., M.Si.

HALAMAN PENGESAHAN

TUGAS AKHIR

ANALISIS KLASIFIKASI WAKTU TUNGGU KERJA DENGAN METODE *SUPPORT VECTOR MACHINE* DAN *NAÏVE BAYES* CLASSIFICATION

(Studi Kasus : Waktu tunggu kerja Alumni Universitas Islam Indonesia)

Nama Mahasiswa : Gustiara Dayu Amrinda

Nomor Mahasiswa : 14 611 163

TUGAS AKHIR INI TELAH DIUJIKAN

PADA TANGGAL

Nama Penguji :	**Tanda Tangan**
1. Ir. Ali Parkhan, M.T	.....
2. Muhammad Muhajir, S.Si, M.Sc	.....
3. Ayundyah Kesumawati, S.Si, M.Si	.....

Mengetahui,

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Drs. Allwar, M.Sc., Ph.D.

HALAMAN PERSEMBAHAN

Tugas akhir ini penulis persembahkan untuk orang-orang yang disayang:

- ✓ ***Ibu Ayundyah Kesumawati, S.Si, M.Si selaku dosen pembimbing skripsi saya, yang sudah membimbing saya dengan sabar. Mengajari saya ilmu yang tidak bisa terbayarkan, mengajari saya artinya berjuang untuk mendapatkan sesuatu yang diinginkan. Terima kasih ibu, jasamu akan saya kenang selalu, Ilmu yang ibu berikan akan saya terapkan seumur hidup saya.***
- ✓ ***Bapak Achmad Darsono dan Ibu Sri Yuneti, selaku Bapak dan Ibu saya yang selalu mendoakan dan berjuang mencari nafkah buat anaknya. Terima kasih untuk semua yang telah diberikan kepada saya walaupun sampai kapanpun saya tidak bisa membalas budi kepada orang tua. Semoga karya ini bisa menjadi hadiah terindah meskipun masih banyak kekurangan.***
- ✓ ***Berlianda Rahma Yuninda, Adik perempuan saya tercinta. Terima kasih ata segala dukungan dan doa yang udah diberika. Semoga Adikku kedepanya semakin sukses dan dipermudah segala urusanya.***
- ✓ ***Sutra Aulia Damayanti, Adik perempuan saya tercinta. Terima kasih atas dukungan dan doanya. Semoga sukses sekolahnya, menjadi kebanggan Bapak dan Ibu.***
- ✓ ***Siapapun dibalik perjuangan saya, sahabat SMP saya, sahabat SMA saya, maupun sahabat Kuliah saya. Terima kasih sudah memberikan cerita dan pengalaman hidup buat saya dan selalu mensupport saya untuk menyelesaikan karya ini.***
- ✓ ***Teman-teman seperjuangan Statistika 2014, terima kasih banyak atas kebersamaan dan dukungannya.***

KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

Assalamu'laikum Wr. Wb.

Alhamdulillah puji syukur kepada Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya, serta shalawat dan salam kepada Nabi Muhammad SAW sehingga penulis dapat menyelesaikan tugas akhir yang berjudul "**Analisis Klasifikasi Waktu Tunggu Kerja dengan Metode *Support Vector Machine* dan *Naïve Bayes Classification* (Studi Kasus : Waktu tunggu kerja Alumni Universitas Islam Indonesia)**" dengan baik.

Tugas akhir ini dilakukan sebagai salah satu persyaratan yang harus dipenuhi dalam menyelesaikan Program Strata Satu di Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Indonesia. Selama mengerjakan dan menyusun tugas akhir ini, penulis banyak memperoleh bantuan dari berbagai pihak, baik berupa saran, kritik, ataupun bimbingan. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada:

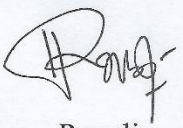
1. Drs. Allwar, M.Sc., Ph.D selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia.
2. Dr. Raden Bagus Fajriya Hakim, S.Si., M.Si selaku Ketua Program Studi Statistika.
3. Ayundyah Kesumawati, S.Si., M.Si selaku dosen pembimbing skripsi atas bimbingan dan kesabarannya selama menyelesaikan tugas akhir ini.
4. Seluruh dosen Statistika di Universitas Islam Indonesia yang telah memberikan ilmunya selama delapan semester.
5. Orang tua tersayang yang selalu setia mendoakan, mendukung, menemani serta memberikan motivasi kepada penulis dalam menyelesaikan tugas akhir ini.
6. Kedua adik kandung yang selalu mendoakan dan mendukung saya hingga selesai penulisan tugas akhir ini.

7. Sahabat saya Treka Heny K.A, Arlina Tri C., Refanina Ulfa R., Meryta Fitri P., dan Winda Pramestiningtyas yang selalu memberikan semangat, mendoakan dan mendukung saya dalam keadaaan apapun.
8. Sahabat saya selama kuliah yaitu 2x+y (Riza Indriani, Erene Fajrila, Nur Hidayah, Inayatus Sholikhah, dan Khoffiya Mulia Rahmi) yang mendengar keluh kesah saya dan ada dalam keadaan apapun.
9. Terima kasih untuk teman belajar kolokium saya dan teman KP saya yang selalu menemani saya dan mengajarkan saya.
10. Terima kasih untuk semua teman prodi statistika angkatan 2014 dan teman statistika lainnya yang tidak bisa penulis sebutkan satu per satu yang selalu menjadi semangat bagi penulis.
11. Pihak lain yang tidak bisa disebutkan satu per satu, terima kasih atas dukungan dan dorongan yang telah diberikan.

Semoga Allah SWT selalu memberikan rahmat dan anugerah-Nya kepada mereka semua atas segala bantuan, bimbingan, dan pengajaran yang telah diberikan kepada penulis. Penulis menyadari sepenuhnya bahwa tugas akhir ini masih jauh dari sempurna. Oleh karena itu penulis menerima kritik dan saran yang bersifat membangun dari semua pihak demi kesempurnaan tugas akhir ini. Semoga tugas akhir ini bermanfaat bagi semua pihak.

Wassalamu'alaikum Wr. Wb.

Yogyakarta, 9 April 2018



Penulis

DAFTAR ISI

HALAMAN JUDUL ..... **i**
HALAMAN PERSETUJUAN PEMBIMBING ..... **iii**
HALAMAN PENGESAHAN ..... **iv**
HALAMAN PERSEMBAHAN ..... **v**
KATA PENGANTAR ..... **vi**
DAFTAR ISI ..... **viii**
DAFTAR TABEL ..... **x**
DAFTAR GAMBAR ..... **xi**
DAFTAR LAMPIRAN ..... **xii**
PERNYATAAN ..... **xiii**
INTISARI ..... **xiv**
ABSTRACT ..... **xv**
BAB I PENDAHULUAN **1**
1.1. Latar Belakang Penelitian 1
1.2. Rumusan Masalah 4
1.3. Batasan Masalah 4
1.4. Jenis Penelitian dan Metode Analisis 5
1.5. Tujuan Penelitian 5
1.6. Manfaat Penelitian 5
BAB II TINJAUAN PUSTAKA **6**
2.1. Waktu Tunggu Kerja Fresh Graduate 6
2.2. Analisis Support Vector Machine (SVM) 7
2.3. Analisis Naïve Bayes Classification (NBC) 8
2.4. Penelitian Penulis 8
BAB III LANDASAN TEORI **9**
3.1. Waktu Tunggu Kerja 9

3.2. *Tracer Study* 9
3.3. Klasifikasi 10
3.4. Data Mining 11
3.5. *Support Vector Machine* 12
3.5.1. Klasifikasi Non-Linear 13
3.6. *Naïve Bayes Classification* 14
3.7. *Accuracy Assessment* 16
3.8. Pengukuran Kinerja Klasifikasi 17
BAB IV METODOLOGI PENELITIAN 18
4.1. Data Penelitian 18
4.3. Variabel dan Definisi Operasional Variabel 18
4.4. Metode Analisis Data 20
4.5. Proses Analisis Data 20
4.6. Tahap Penelitian 21
BAB V HASIL DAN PEMBAHASAN 23
5.1. Analisis Deskriptif 23
5.2. Klasifikasi menggunakan *Support Vector Machine* (SVM) 30
5.2.1. Perhitungan *Support Vector Machine* 34
5.3. Klasifikasi Menggunakan Naïve Bayes Classification (NBC) 38
5.4. Perbandingan Analisis SVM dan NBC 41
BAB VI PENUTUP 42
6.1. Kesimpulan 42
6.2. Saran 42
DAFTAR PUSTAKA 43
LAMPIRAN .....

DAFTAR TABEL

Nomor Judul Halaman

Tabel 3.1 Macam Fungsi Kernel 12
Tabel 3.2 *Confussion Matrix* Tiga Kelas 16
Tabel 4.1 Definisi Variabel 18
Tabel 4.3 Variabel Y 2 Kelas 21
Tabel 5.1 Tabel Analisis Deskriptif 29
Tabel 5.2 Akurasi Data *Training* SVM 31
Tabel 5.3 Prediksi *Training* SVM 31
Tabel 5.4 Hasil Akurasi *Training* SVM 32
Tabel 5.5 Akurasi Data *Testing* SVM 32
Tabel 5.6 Hasil Prediksi Polynomial *Testing* SVM 33
Tabel 5.7 Hasil Akurasi *Testing* SVM 33
Tabel 5.8 Menghitung Matrik Kernel 35
Tabel 5.9 Menghitung Menggunakan Data Uji 37
Tabel 5.10 A-Priori Probability 38
Tabel 5.11 Hasil Prediksi *Training* NBC 39
Tabel 5.12 Hasil Akurasi *Training* NBC 39
Tabel 5.13 Hasil Prediksi *Testing* NBC 40
Tabel 5.14 Hasil Akurasi *Testing* NBC 40
Tabel 5.15 Perbandingan Klasifikasi SVM dan NBC 41

DAFTAR GAMBAR

Nomor Judul Halaman

Gambar 4.1 Diagram Alir Tahap Penelitian 22
Gambar 5.1 Presentase Jenjang Studi 23
Gambar 5.2 Presentase Fakultas 24
Gambar 5.3 Presentase Jurusan 24
Gambar 5.4 Presentase Indeks Prestasi Kumulatif (IPK) 25
Gambar 5.5 Presentase Jenis Kelamin 26
Gambar 5.6 Presentase Ketepatan Studi 27
Gambar 5.7 Presentase Organisasi 28
Gambar 5.8 Waktu Tunggu Kerja 2 Kelas 29
Gambar 5.9 *Hyperplane* 38

DAFTAR LAMPIRAN

Lampiran 1 : *Crosstabulation* Data

Lampiran 2 : Hasil analisis menggunakan *software R*

PERNYATAAN

Dengan ini penulis menyatakan bahwa dalam tugas akhir ini tidak terdapat karya yang sebelumnya pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi dan sepengetahuan penulis juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 9 April 2018



Penulis

ANALISIS KLASIFIKASI WAKTU TUNGGU KERJA DENGAN METODE *SUPPORT VECTOR MACHINE* DAN *NAÏVE BAYES CLASSIFICATION*

INTISARI

Perguruan tinggi merupakan tingkat pendidikan yang mampu menghasilkan dan mempersiapkan lulusan maka lulusan perguruan tinggi mampu bersaing dan siap menghadapi dunia kerja. Peneliti ingin mengklasifikasikan waktu tunggu kerja lulusan tersebut apakah termasuk lancar dan tidak lancar untuk mengetahui seberapa tepat waktu tunggu kerja dapat diklasifikasikan dengan baik. Tujuan dari penelitian ini adalah untuk mengetahui seberapa tepat dengan menggunakan metode *Support Vector Machine* (SVM) dan *Naïve Bayes Classification* (NBC) dalam mengklasifikasikan waktu tunggu kerja lulusan mahasiswa UII. Data yang digunakan adalah data sekunder yang diperoleh dari UII Career Center yaitu data *tracer study* alumni UII pada tahun 2015. Hasil analisis menggunakan *Support Vector Machine* (SVM) dan *Naïve Bayes Classification* (NBC) menunjukkan bahwa akurasi dari kedua metode yang paling baik hasil klasifikasi 2 kelas dengan nilai akurasinya yang paling tinggi adalah SVM menggunakan nilai *cost* 0,1 dan *degree* 1 sehingga diperoleh klasifikasi yang baik sebesar 0,6875 atau 68,75%.

Kata Kunci: Waktu Tunggu Kerja, *Tracer Study*, SVM, NBC.

ANALYSIS OF WAITING TIME OF WORK WITH THE CLASSIFICATION METHOD OF SUPPORT VECTOR MACHINE AND NAÏVE BAYES CLASSIFICATION

ABSTRACT

The College is an educational level which is able to produce and prepare graduates then college graduates are able to compete and ready to face the world of work. Researchers like to classify these graduates work waiting time does include smoothly and not smoothly to know how precisely the waiting time can be classified with the good work. The purpose of this research is to find out how exactly to use the method of Support Vector Machine (SVM) and Naïve Bayes Classification (NBC) in classify the work waiting time graduate student UII. The Data used is data secondary obtained from UII Career Center i.e. data tracer study alumni UII in the year 2015. The results of the analysis using Support Vector Machine (SVM) and Naïve Bayes Classification (NBC) showed that the accuracy of the two methods is the best the results of the classification of class 2 with a value of fit is the most high is SVM uses the value of the cost and 0.1 degree 1 so obtained a good classification or 0.6875 of 68.75%.

***Keywords*:** *Waiting Time Of Work, Tracer Study*, SVM, NBC.

BAB I

PENDAHULUAN

1.1. Latar Belakang

Di masa sekarang pendidikan sangatlah penting seiring dengan kemajuan teknologi yang menuntut manusia untuk memiliki pengetahuan luas dan berkualitas. Salah satu cara membentuk manusia berkualitas yaitu berdasarkan tingkat pendidikan yang semakin tinggi. Perguruan tinggi merupakan tingkat pendidikan yang mampu mencetak dan mempersiapkan lulusan–lulusan maka lulusan perguruan tinggi mampu bersaing dan siap menghadapi dunia kerja. Kemampuan lulusan dalam bersaing di dunia pekerjaan menjadi ukuran suatu perguruan tinggi untuk mencetak lulusan yang berkualitas.

Lulusan yang bekerja merupakan tolak ukur suatu perguruan tinggi dapat mencetak dan menjadikan suatu keberhasilan dalam bidang pendidikan sehingga sangat penting. Setiap perguruan tinggi memiliki cara dan sistem tersendiri untuk membentuk lulusan berkualitas. Ristekdikti telah menetapkan standar penilaian yang sama untuk setiap perguruan tinggi. Menurut Direktorat Jenderal Pembelajaran dan Kemahasiswaan (Ristekdikti, 2016) Perguruan tinggi agar melakukan *Tracer Study* secara reguler setiap tahun untuk memenuhi kebutuhan data akreditasi, pengembangan kurikulum dan perbaikan pembelajaran di perguruan tinggi tidak bermaksud menyeragamkan atau mengurangi otonomi perguruan tinggi, kompilasi data Nasional yang akurat akan sangat dibutuhkan untuk pengambilan kebijakan terkait pendidikan tinggi di tingkat Nasional dengan menggunakan data *Tracer Study* yang juga disosialisasikan.

Tracer Study dapat mengukur dan melacak kinerja dan saran-saran lulusan maupun pengguna lulusan sehingga dapat diperoleh indikator yang jelas tentang

jumlah, profil kerja masa mendatang, serta kompetensi yang diperlukan oleh dunia kerja. Berbekal hasil penelitian ini dapat digunakan untuk membenahi program pembelajaran agar mampu membangun kompetensi mahasiswa sesuai dengan kompetensi yang diperlukan oleh dunia kerja. Giliran berikutnya, dengan kompetensi yang memadai, maka diharapkan dapat meningkatkan daya saing lulusan. (Dwihartanti, 2015)

Di Indonesia Perguruan tinggi tersebar hingga ke seluruh wilayah. Salah satunya yaitu di kota Yogyakarta. Yogyakarta merupakan kota pelajar, dimana terdapat banyak universitas didalamnya. Salah satu perguruan tinggi swasta terbaik di Indonesia yaitu Universitas Islam Indonesia yang letaknya berada di wilayah Sleman, Yogyakarta. Universitas Islam Indonesia menjadi perguruan tinggi swasta terbaik dikarenakan lulus penilaian oleh Badan Akreditasi Nasional Perguruan Tinggi dengan nilai institusi yaitu A. Salah satu aspek penilaiannya yaitu dilihat berdasarkan standar Profil lulusan dimana banyaknya mahasiswa yang menyelesaikan studi selama batas masa studi dan Pelacakan serta perekaman data lulusan seperti waktu tunggu kerja pertama dan sebagainya. Lama waktu tunggu kerja lulusan kemungkinan dapat dipengaruhi banyak faktor. Faktor-faktor yang memungkinkan dapat mempengaruhi lama waktu tunggu kerja yaitu Fakultas, Indeks Prestasi Kelulusan (IPK), jenis kelamin, jenjang pendidikan, ketepatan studi, dan organisasi. Artinya informasi dari lulusan perguruan tinggi mengindikasikan sebagai bahan pertimbangan dalam penyusunan standar kompetensi. Banyak mahasiswa yang lulus kemudian menunggu pekerjaan hingga bertahun-tahun, ada juga yang mempeloreh pekerjaan setelah lulus ini dapat mempengaruhi penilaian perguruan tinggi bagaimana perguruan tinggi dapat mencetak lulusannya dalam memperoleh pekerjaan. Alasan peneliti memilih variabel tersebut yaitu karena peneliti mengacu pada penelitian sebelumnya selain itu ketersediaan data *tracer study* yang dimiliki UII Career Center yang sesuai dengan penelitian sebelumnya sehingga peneliti hanya mengambil variabel tersebut.

Menurut Kepala Divisi Penelusuran & Pemberdayaan Alumni (Darmawan, 2017) pada Diseminasi Hasil Tracer Studi, "responden yang ikut berpartisipasi didominasi oleh alumni yang Bekerja sebanyak 59,36%, sebanyak 20,53% dari total responden mendapatkan pekerjaan dengan masa tunggu 1 bulan, yang mana merupakan jumlah terbanyak berdasarkan masa tunggu dalam bulan. Selain itu, sebanyak 15,40% lainnya memiliki masa tunggu 0 bulan, dengan kata lain responden telah memperoleh pekerjaan sebelum masa kelulusan. Sementara responden yang menunggu selama 2 bulan dan 3 bulan masing-masing sebanyak 10,35% dan 12,39%. Total responden yang memiliki masa tunggu kurang dari sama dengan 3 bulan adalah sebanyak 58,67%." Maka sisanya adalah masa tunggu >3 bulan. Baziedy menambahkan "telah terjadi peningkatan pengisian *tracer study* selama 1 tahun terakhir. Di tahun 2017, UII berhasil mendapatkan lebih dari 2.400 responden yang berpartisipasi dalam *tracer study* UII sedangkan pada tahun 2016 sendiri ada 1.906 responden". Artinya dimana pada data tersebut tidak semua alumni yang bekerja mengembalikan kuisioner tersebut data yang dibutuhkan oleh UII.

Pada penelitian ini data yang digunakan adalah data yang berasal dari UII Career Center yaitu data *Tracer Study* tahun 2015. Di UII sendiri memiliki standar berkarya setelah lulus yaitu maksimal 3 bulan karena belum pernah dilakukan penelitian terkait pengklasifikasian waktu tunggu kerja sehingga peneliti ingin mengklasifikasikan waktu tunggu kerja dan juga peneliti mencoba menerapkan kategori waktu tunggu kerja seperti penelitian sebelumnya. Berdasarkan permasalahan tersebut maka peneliti ingin mengklasifikasikan waktu tunggu kerja lulusan dengan menggunakan metode *Support Vector Machine* (SVM) dan *Naïve Bayes Classification* (NBC). Pengklasifikasian data menggunakan metode SVM dikarenakan memiliki tingkat akurasi yang paling baik ini dibuktikan dengan penelitian yang dilakukan oleh Fakhriyani (2016) yaitu Perbandingan Algoritma *Naïve Bayes Classification* Dan *Support Vector Machine* Dalam Seleksi Kelulusan Pemberkasan Beasiswa BPP-PPA Fakultas Teknik Universitas Negeri Jakarta diperoleh kesimpulan bahwa akurasi

algoritma *Naïve Bayes Classification* adalah 0.7542, sedangkan hasil akurasi algoritma *Support Vector Machine* adalah 0.99.

Penulis tertarik untuk menggunakan klasifikasi pada waktu tunggu kerja lulusan Universitas Islam Indonesia di Yogyakarta dengan metode *Support Vector Machine* (SVM) karena metode ini mempunyai tingkat akurasi terbaik dan metode *Naïve Bayes Classification* (NBC) karena metode ini merupakan metode yang paling sederhana. Penelitian ini bertujuan untuk mengetahui seberapa tepat akurasi yang diperoleh dari metode klasifikasi *Support Vector Machine* (SVM) dan *Naïve Bayes Classification* pada data waktu tunggu kerja alumni Universitas Islam Indonesia di Yogyakarta.

1.2. Rumusan Masalah

Berdasarkan latar belakang permasalahan di atas maka rumusan masalah yang dapat diidentifikasi dalam tugas akhir ini adalah:

1. Bagaimana gambaran dari faktor yang mendukung waktu tunggu lulusan mahasiswa UII?
2. Seberapa tepat dengan menggunakan metode *Support Vector Machine (SVM)* dan *Naïve Bayes Classification* (NBC) dalam mengklasifikasikan waktu tunggu kerja lulusan mahasiswa UII?

1.3. Batasan Masalah

Batasan masalah sangat diperlukan agar tidak terjadi penyimpangan. Untuk itu batasan masalah dalam penulisan ini adalah sebagai berikut:

1. Penelitian ini dilakukan pada data yang diambil dari UII Career Center dimana data *tracer study* alumni UII pada tahun 2015.
2. Alat analisis yang digunakan untuk analisis waktu tunggu kerja lulusan adalah deskriptif dan metode SVM dan NBC
3. Waktu tunggu kerja diklasifikasikan menjadi 2 yaitu lancar dan tidak lancar.
4. Peneliti menentukan standar akurasi terbaik yaitu akurasi lebih dari 50% dapat dikatakan baik.

1.4. Jenis Penelitian dan Metode Analisis

Tugas akhir ini termasuk dalam kategori aplikasi. Metode analisis yang digunakan adalah Analisis Deskriptif, *Support Vector Machine* (SVM) dan *Naïve Bayes Classification* (NBC).

1.5. Tujuan Penelitian

1. Mengetahui bagaimana gambaran umum dari faktor yang mendukung waktu tunggu lulusan mahasiswa UII.
2. Mengetahui seberapa tepat dengan menggunakan metode *Support Vector Machine* (SVM) dan *Naïve Bayes Classification* (NBC) dalam mengklasifikasikan waktu tunggu kerja lulusan mahasiswa UII serta mengetahui kelas terbaik.

1.6. Manfaat Penelitian

Dengan diperolehnya hasil dari metode analisis deskriptif sebagai gambaran faktor dari waktu tunggu kerja lulusan mahasiswa UII dan dengan metode SVM dan NBC dapat mengetahui ketepatan metode yang digunakan dan diperoleh akurasi terbaik dalam mengklasifikasikan waktu tunggu kerja lulusan UII.

BAB II
TINJAUAN PUSTAKA

Hal yang paling penting dalam membuat suatu penelitian adalah penelitian terdahulu. Penelitian yang dilakukan saat ini mengacu pada penelitian yang sebelumnya untuk mengetahui hubungan antara penelitian sebelumnya dan penelitian terdahulu digunakan sebagai kajian bagi penulis. Dengan menggunakan penelitian terdahulu dapat menghindari unsur duplikasi penelitian, selain itu penelitian ini mendapat acuan dan kontribusi pada perkembangan pengetahuan untuk penelitian selanjutnya.

2.1 Waktu Tunggu Kerja *Fresh Graduate*

Nama Peneliti (Mashuri, 2012) Variabel Penelitian yang digunakan ialah Jenis kelamin, Tempat tinggal asal, Status SLTA, Jumlah anggota keluarga, Kondisi alumni selama kuliah sampai dengan kuliah terdiri dari tempat tinggal selama kuliah, Pegeluaran uang perbulan selama kuliah, Penempatan kerja/magang, Pengalaman kerja selama kuliah, Pengalaman organisasi, Pengalaman dalam kepanitiaan, Pengalaman mengikuti kompetisi/perlombaan, Beasiswa, pengalaman mengikuti pelatihan, waktu kelulusan, jenis bidang yang diminati, partisipasi mahasiswa dalam penelitian dengan dosen, nilai IPK, dan nilai TOEFL. Sedangkan variabel lainnya ada cara mencari pekerjaan, jenis pekerjaan yang dipilih, dan motivasi dalam mencari kerja. Metode yang digunakan pada penelitian ini yaitu analisis Regresi Logistik Ordinal. Hasil dari penelitian yaitu variabel yang berpengaruh terhadap waktu tunggu kerja alumni salah satunya adalah variabel jenis pekerjaan yang dipilih alumni setelah lulus. Berdasarkan model regresi logistik ordinal yang terbentuk adalah sebesar 33,93% dan ketepatan klasifikasi sebesar 66,07%.

Nama Peneliti (Fitri, 2010) Variabel Penelitian yang digunakan ialah enam kekuatan Motivasi dan Empat Tipe Motivasi dengan lamanya masa tunggu kerja. Metode yang digunakan pada penelitian ini yaitu analisis uji korelasi Spearman dan Regresi. Hasil dari penelitian yaitu Hasil analisa menggunakan uji korelasi *Spearman* diperoleh semakin tinggi motivasi seseorang maka semakin singkat masa tunggu kerja dan sebaliknya, semakin rendah motivasi seseorang maka semakin lama rentang waktu yang dibutuhkan untuk memperoleh pekerjaan. Uji regresi didapatkan hasil,dari enam kekuatan motivasi kekuatan demokratik dan teknokrat merupakan kekuatan yang memberikan konteribusi paling besar dalam menentukan tinggi rendahnya motivasi lulusan.

2.2 Analisis *Support Vector Machine (SVM)*

Nama Peneliti (Hoyyi, 2016) Variabel Penelitian yang digunakan ialah lama studi, jenis kelamin, IPK, Beasiswa, *Part time* Organisasi, dan Jalur Masuk UNDIP. Metode yang digunakan pada penelitian ini yaitu analisis *Support Vector Machine* dan *Iterative Dichotomiser 3*. Hasil dari penelitian ini yaitu perbandingan antara kedua metode pada data *training* memperoleh hasil yang baik sebesar 90% namun pada data *training* metode ID3 lebih baik dibandingkan SVM.

Nama Peneliti (Pusphita Anna Octaviani, 2014). Variabel Penelitian yang digunakan ialah status akreditasi, standar isi, standar proses, standar kompetensi lulusan, standar pendidik dan tenaga pendidikan, standar sarana dan prasarana, standar pengelolaan, dan standar penilaian pendidikan. Metode yang digunakan pada penelitian ini yaitu analisis *Support Vector Machine*. Hasil dari penelitian ini yaitu pada klasifikasi akreditasi sekolah dasar dengan data *training* 337 data menggunakan fungsi kernel Gaussian *Radial Basic Function* (RBF) sebesar 100% sedangkan akurasi menggunakan fungsi kernel polynomial sebesar 98.810%. sedangkan klasifikasi akreditasi sekolah dasar dengan data *testing* 82 data menggunakan fungsi kernel Gaussian *Radial Basic Function* (RBF) sebesar 93.902% sedangkan akurasi menggunakan fungsi kernel polynomial sebesar 92.683%

2.3 Analisis *Naïve Bayes Classification (NBC)*

Nama peneliti (Faramudhita, 2018) Variabel Penelitian yang digunakan ialah IPK, Jenis Kelamin, Kota Lahir, Tipe Sekolah, Kota Sekolah, dan Lama Studi. Metode yang digunakan pada penelitian ini yaitu Analisis Regresi Logistik Ordinal dan klasifikasi Naïve Bayes pada data alumni UNILA tahun 2016. Hasil dari penelitian ini yaitu dalam memprediksi tingkat eror pada klasifikasi yang digunakan berkisar pada 20% sampai 34% yang hal ini bisa dipengaruhi berdasarkan data *training* ataupun *testing* yang digunakan serta kekonsisten data yang digunakan.

Nama peneliti (Bhawiyuga, 2018) Variabel yang digunakan ialah app Name, Src TCP Flags Description, total Source Bytes, Dst TCP Flags Description, total Destination Bytes, Source, total Destination Packet, protocol Name, total Source Packet, source Port, Src Payload As Base 64, Destination, Src Payload As Base UTF, destination Port, Dst Payload As Base64, start Date Time, Dst Payload As Base UTF, stop Date Time, dan direction. Metode yang digunakan pada penelitian ini yaitu Analisis *Naïve Bayes* dan *Support Vector Machine* (SVM). Hasil dari penelitian ini yaitu dalam hasil perolehan nilai dari *confusion matrix* pada masing-masing kelasifier *Naive Bayes Classification* sebesar 85.055%, SVM Linear sebesar 99.995%, SVM Polynomial sebesar 99.999%, dan SVM Sigmoid sebesar 99.995% menghasilkan persentase akurasi yang baik.

2.4 Penelitian Penulis

Berdasarkan beberapa referensi dari penelitian terdahulu, peneliti melakukan penelitian klasifikasi waktu tunggu kerja lulusan Universitas Islam Indonesia di Yogyakarta dengan metode *Support Vector Machine* (SVM) dan *Naïve Bayes Classification*. Variabel yang digunakan yaitu Jenjang Studi, Fakultas, Indeks Prestasi Kumulatif, Jenis Kelamin, Masa Studi, Organisasi, dan Waktu Tunggu Bekerja.

BAB III

LANDASAN TEORI

3.1 Waktu Tunggu Kerja

Menurut (Muslim, 2009), Kerja merupakan suatu perbuatan yang dilakukan, diperbuat atau mencari. Sedangkan pekerjaan merupakan suatu yang dilakukan untuk mencari nafkah atau hal mengerjakan sesuatu. "waktu tunggu" terdiri dari dua kata yaitu kata "waktu" dan kata "tunggu" keduanya memiliki arti tersendiri. Kata "waktu" berasal dari kata sifat yaitu waktu yang panjang. Sedangkan kata "tunggu" berasal dari kata yaitu tinggal sementara untuk berjaga atau menanti. Jadi dapat disimpulkan waktu tunggu kerja adalah masa dimana lulusan menunggu untuk mendapat pekerjaan dari awal lulus kuliah hingga mendapat pekerjaan. Pada dasarnya kalimat waktu tunggu kerja disini dapat diartikan, disebut juga dengan "masa menganggur", karena dalam jangka menunggu kerja tersebut seseorang akan menjadi seorang pengangguran. Hakekat pengangguran adalah suatu keadaan dimana orang yang ingin dan mampu bekerja, tidak dapat memperoleh pekerjaan.

3.2 *Tracer Study* UII

Tracer Study UII adalah survei mengenai lulusan dalam rangka menyediakan berbagai informasi yang bermanfaat bagi kepentingan evaluasi hasil dari proses pendidikan tinggi yang diselenggarakan oleh UII. Tujuan dari *Tracer Study* ini untuk menggali berbagai informasi yang berasal dari alumni untuk dimanfaatkan bagi kepentingan evaluasi UII dan melakukan penyempurnaan sekaligus penjaminan kualitas UII sebagai lembaga pendidikan tinggi.

Manfaat *Tracer Study* yaitu dalam menyediakan informasi-informasi penting mengenai hubungan antara perguruan tinggi dan dunia kerja, menilai relevansi pendidikan tinggi dengan dunia industri, serta memberikan informasi bagi para

pemangku kepentingan (Alumni, Civitas Akademika UII, Dunia Industri, Mahasiswa UII, Orang Tua Mahasiswa, Calon Mahasiswa, dsb.). Selain itu juga sebagai kelengkapan persyaratan dan pendukung bagi proses Akreditasi UII, baik Akreditasi Nasional maupun Akreditasi Internasional.

Pelaksanaa *Tracer Study* UII pada tahun 2017 ini yaitu dengan pengisian kuesioner *Tracer Study* dari para alumni UII pada program D3, S1, dan Program Profesi yang lulus pada tahun 2015. Teknik pengambilan data menggunakan *e-mail blast, SMS blast,* dan *phone call*. Pengisian kuesioner *Tracer Study* oleh para alumni dilakukan pada laman *web* tracer.uii.ac.id. Alumni dapat mengisi kuesioner *Tracer Study* dengan log in menggunakan nomor mahasiswa (saat menempuh perkuliahan D3, S1, atau Program Profesi di UII dan lulus pada tahun 2015).

Alumni sangat berpartisipasi dalam mengisi kuesioner *Tracer Study* ini akan memberikan masukan bagi pengembangan UII di masa yang akan datang. Selain itu hasil kuisioner ini sangat berkontribusi bagi peningkatan dan penjaminan kualitas pendidikan tinggi di UII dan sekaligus mendukung proses Akreditasi Nasional maupun Akreditasi Internasional UII di masa yang akan datang. (Center, 2017)

3.3 Klasifikasi

Pada Data mining kasus klasifikasi ada 2 yaitu bersifat linear dan tidak linear. Kelinearan pada data yang diasumsikan menggunakan banyak metode yang dikembangkan. Metode itu pun memiliki performansi yang kurang baik ketika diterapkan pada kasus-kasus nonlinear. (Santosa, 2007). Metode analisis data yang digunakan untuk membentuk model mendeskripsikan kelas data penting atau model mendeskripsikan *trend* data adalah klasifikasi. Fungsinya untuk memprediksi kelas data yang bersifat *categorical*, sedangkan prediksi untuk memodelkan fungsi yang bernilai *continuous*. (Furey, 2000)

3.4 Data Mining

Menurut (Han J., 2000) penyebab utama data mining diperlukan karena adanya sejumlah besar data yang dapat digunakan untuk menghasilkan informasi dan pengetahuan yang berguna. Informasi dan pengetahuan yang diperoleh tersebut dapat digunakan pada banyak bidang, mulai manajemen bisnis, kontrol produksi, kesehatan, dan lain-lain. Berikut ini tahapan Data *Mining*:

1. Pembersihan data (Data *Cleaning*)

   Pembersihan data adalah proses menghilangkan *noise* dan data tidak relevan atau data yang tidak konsisten.
2. Integrasi data (Data *Integration*)

   Integrasi data ialah penggabungan data dari berbagai *database* ke dalam satu *database* baru.
3. Seleksi data (Data *Selection*)

   Tidak semua data yang ada pada database sering kali digunakan, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database.
4. Transformasi data (Data *Transformation*)

   Data diubah atau digabung ke dalam format yang sesuai supaya diproses dalam Data *Mining*.
5. Proses *Mining*

   Suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga yang tersembunyi dari data. Beberapa metode yang dapat digunakan berdasarkan pengelompokan Data *Mining*.
6. Evaluasi pola (*Pattern Evaluation*)

   Untuk mengidentifikasi pola-pola menarik ke dalam pengetahuan dasar yang ditemukan.
7. Presentasi pengetahuan (*Knowledge Presentation*)

   ialah visualisasi dan penyajian pengetahuan tentang metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. (Ridwan, 2013).

3.5 *Support Vector Machine*

*Support Vector Machine (*SVM) merupakan metode mesin pembelajaran (*learning machine*) dengan tujuan menemukan fungsi pemisah (*hyperplane*) terbaik yang memisahkan dua buah kelas pada *input space* (Nugroho dkk, 2003). Konsep SVM adalah untuk mencari fungsi pemisah terbaik fungsinya sebagai pemisah antara dua buah kelas pada *input space*. (Cristanini dan Jhon, 2000).

SVM merupakan salah satu metode terbaik yang bisa dipakai dalam permasalahan klasifikasi. Konsep SVM berawal dari masalah klasifikasi dua kelas sehingga membutuhkan *training set* positif dan negatif. Pada beberapa kasus, data tidak bisa diklasifikasi menggunakan metode *SVM Linear*, sehingga dikembangkan fungsi *kernel* untuk mengklasifikasikan data dalam bentuk non-*linear*.

Beberapa Macam Fungsi Kernel *Support Vector Machine* (SVM) :

Tabel 3.1 Macam Fungsi Kernel

No	Nama Kernel	Definisi Fungsi
1	Linear	$K(x,y) = x.y$
2	Polinomial of degree d	$K(x,y) = (x.y)^d$
3	Polinomial of degree up to d	$K(x,y) = (x.y + c)^d$
4	Gaussian RBF	$K(x,y) = exp\left(\frac{-\|x-y\|^2}{2\sigma^2}\right)$
5	Sigmoid (Tangen Hiperbolik)	$K(x,y) = tanh(\sigma(x.y) + c)$
6	Invers Multi Kuadratik	$K(x,y) = \frac{1}{\sqrt{\|x-y\|^2+c}}$
7	Additive	$K(x,y) = \sum_{i=1}^{n} K(x_i, y_i)$

Kernel Linear diperlukan ketika data yang akan diklasifikasi dapat terpisah dengan sebuah garis/*hyperplane*. Kernel non-Linear diperlukan ketika data hanya dapat dipisahkan dengan garis lengkung atau sebuah bidang pada ruang dimensi tinggi (Kernel Trick, No.2 sampai 6). Klasifikasi dua kelas pengembangan dari SVM yang dikembangkan lagi menjadi klasifikasi multikelas (Santosa, 2007) dimana *hyperplane* lebih dari satu yang terbentuk.

3.5.1 Klasifikasi Non-Linear

Masalah dalam dunia nyata (*real world problem*) pada umumnya jarang yang bersifat linear *separable* (tidak terpisahkan secara linear), tetapi bersifat non-linear (Nugroho dkk, 2003). Menurut (Prasetyo, 2012) Pada problem non-linear untuk menyelesaikan, SVM dimodifikasi dengan memasukkan fungsi kernel. Kernel dapat diartikan sebagai fungsi yang memetakan fitur data dari dimensi awal (rendah) ke fitur yang lebih tinggi (bahkan jauh lebih tinggi). Metode klasifikasi pada umumnya menggunakan pendekatan berbeda yang justru mengurangi dimensi awal untuk menyederhanakan proses komputasi dan menghasilkan akurasi prediksi yang lebih baik. Misalnya untuk *n* sampel data $\big((\Phi(x_1), y_1); (\Phi(x_2), y_2); \dots; (\Phi(x_n), y_n)\big)$, *dot product* dua buah vektor ($\boldsymbol{x}_i$) dan ($\boldsymbol{x}_j$) dinotasikan sebagai. $\Phi(x_i).\Phi(x_t)$. Tanpa mengetahui fungsi transformasi Φ, nilai *dot product* tersebut dapat dihitung menggunakan komponen kedua buah vektor tersebut di ruang dimensi asal, seperti berikut:

$$K(x_i, x_t) = \Phi(x_i).\Phi(x_t) \tag{3.1}$$

Dimana:

$\Phi(x_i).\Phi(x_t)$: *dot product* dua buah vektor ($\boldsymbol{x}_i$) dan ($\boldsymbol{x}_j$)

Nilai $K(x_i, x_t)$ adalah fungsi kernel yang menunjukkan pemetaan non-linear pada *feature space*. Prediksi sekumpulan data dengan dimensi fitur yang baru dengan persamaan berikut:

$$f\big(\Phi(x)\big) = \text{sign}(w.\Phi(x_t) + b) = \text{sign}(\textstyle\sum_{i=1}^{ns} \alpha_i y_i \Phi(x_i).\Phi(x_t) + b) \tag{3.2}$$

$$= \text{sign}(\textstyle\sum_{i=1}^{ns} \alpha_i y_i K(x_i, x_t) + b) \tag{3.3}$$

Dimana:

ns : jumlah data menjadi *support vector*

x_i : *support vector*

x_t : data *testing* yang akan diprediksi

klasifikasi menggunakan Fungsi Kernel terdapat beberapa nilai C (*cost*) yang digunakan untuk menentukan ketepatan klasifikasi, yaitu 0.1, 1, 10, dan 100 dan parameter d(*degree*) berbeda untuk tiap jenis kernel.

3.6 *Naïve Bayes Classification*

Naïve Bayes Classification adalah sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. *Naïve Bayes Classification* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya.

Naïve Bayes Classification didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai *output*. Keuntungan penggunaan *Naïve Bayes Classification* yaitu metode ini hanya membutuhkan jumlah data pelatihan (*Training* Data) yang kecil untuk menentukan estimasi paremeter yang diperlukan dalam proses pengklasifikasian. *Naïve Bayes Classification* sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan.

Persamaan Metode *Naïve Bayes Classification*. Persamaan dari teorema *Bayes* adalah:

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \tag{3.4}$$

Di mana:

X	: Data dengan *kelas* yang belum diketahui
H	: Hipotesis data merupakan suatu *kelas* spesifik
P(H|X)	: Probabilitas hipotesis *H* berdasar kondisi *X* (posteriori probabilitas)
P(H)	: Probabilitas hipotesis *H* (prior probabilitas)
P(X|H)	:Probabilitas *X* berdasarkan kondisi pada hipotesis *H*
P(X)	: Probabilitas *X*

Untuk menjelaskan metode *Naive Bayes Classification*, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, metode *Naive Bayes Classification* di atas disesuaikan sebagai berikut:

$$P(C|F1 \dots Fn) = \frac{P(C)P(F1 \dots Fn|C)}{P(F1 \dots Fn)} \quad (3.5)$$

Di mana Variabel *C* merepresentasikan kelas, sementara variabel *F1 ... Fn* merepresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas *C* (*Posterior*) adalah peluang munculnya kelas *C* (sebelum masuknya sampel tersebut, seringkali disebut *prior*), dikali dengan peluang kemunculan karakteristik-karakteristik sampel pada kelas *C* (disebut juga *likelihood*), dibagi dengan peluang kemunculan karakteristik-karakteristik sampel secara global (disebut juga *evidence*). Karena itu, rumus di atas dapat pula ditulis secara sederhana sebagai berikut:

$$Posterior = \frac{Prior\ x\ likelihood}{evidence} \quad (3.6)$$

Nilai *Evidence* selalu tetap untuk setiap kelas pada satu sampel. Nilai dari *posterior* tersebut nantinya akan dibandingkan dengan nilai-nilai *posterior* kelas lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan. Penjabaran lebih lanjut rumus *Bayes* tersebut dilakukan dengan menjabarkan ($C|F1,\dots,$) menggunakan aturan perkalian sebagai berikut:

$$P(C|F1, \dots, Fn = P(C)P(F1, \dots, Fn|C) \ = P(C)P(F1|C)P(F2, \dots, Fn|C, F1)$$
$$= P(C)P(F1|C)P(F2|C, F1)P(F3, \dots, Fn|C, F1, F2$$
$$= (C)P(F1|C)P(F2|C, F1)P(F3|C, F1, F2)P(F4, \dots, Fn|C, F1, F2, F3)$$
$$= P(C)P(F1|C)P(F2|C, F1)P(F3|C, F1, F2) \dots P(Fn|C, F1, F2, F3, \dots, Fn - 1) \quad (3.7)$$

Hal ini menunjukkan bahwa hasil penjabaran tersebut menyebabkan semakin banyak dan semakin kompleksnya faktor-faktor yang mempengaruhi nilai probabilitas, yang hampir mustahil untuk di analisa satu persatu. Menyebabkan perhitungan tersebut

sulit untuk dilakukan. Sehingga diperlukan asumsi indepensi yang sangat tinggi (naif), masing-masing petunjuk (F1,F2..Fn) saling bebas satu sama lain. Berdasarkan asumsi tersebut berlaku persamaan berikut:

$$P(F_i|F_j) = \frac{P(F_i \cap F_j)}{P(F_j)} = \frac{P(F_i)P(F_j)}{P(F_j)} = P(F_i) \quad (3.8)$$

Dimana i≠j, sehingga

$$P(F_i|C, F_j) = P(F_i|C) \quad (3.9)$$

Persamaan menunjukkan model dari teorema *Naïve bayes Classification* yang selanjutnya digunakan untuk proses klasifikasi. (Saleh, 2015)

3.7 *Accuracy Assessment*

Proses klasifikasi belum sempurna apabila belum menghitung tingkat keakuratannya (Lillesand, 2005). Pada suatu penelitian klasifikasi dikenal dengan istilah *The Kappa Index of Agreement* (KIA). Nilai kappa adalah ukuran kebenaran antara kelas yang direpresentasikan di dalam waktu tunggu kerja. Kappa merupakan nilai kecocokan hasil klasifikasi waktu tunggu kerja dengan keadaan yang sebenarnya dilapangan. Semakin tinggi nilai Kappa maka semakin benar klasifikasi yang dilakukan, sedangkan semakin rendah nilai Kappa maka semakin rendah kemungkin hasil klasifikasinya dapat digunakan.

3.8 Pengukuran Kinerja Klasifikasi

Pada model klasifikasi data asli dan data hasil dibentuk tabulasi silang (matriks konfusi) berisi informasi terkait kelas data asli yang disajikan dalam baris matriks dan kelas data hasil prediksi disajikan pada kolom klasifikasi (Prasetyo,2012).

Tabel 3.2 Confussion Matrix tiga kelas

F_{gh}		Kelas Prediksi (h)		
		Kelas 1	Kelas 2	Kelas 3
Kelas Asli (g)	Kelas 1	F_{11}	F_{12}	F_{13}
	Kelas 2	F_{21}	F_{22}	F_{23}
	Kelas 3	F_{31}	F_{32}	F_{33}

Menurut (Prasetyo, 2012) ketepatan akurasi dilihat berdasarkan nilai akurasi klasifikasi. Nilai pada akurasi klasifikasi menunjukkan performasi model dari klasifikasi secara menyeluruh, sehingga semakin tinggi nilai akurasi klasifikasi maka semakin baik pula performasi dari model klasifikasi.

$$\text{Akurasi Klasifikasi} = \frac{F_{11}+F_{22}+F_{33}}{F_{11}+F_{12}+F_{13}+F_{21}+F_{22}+F_{23}+F_{31}+F_{32}+F_{33}} \quad (3.10)$$

$$\text{Error Klasifikasi} = \frac{F_{12}+F_{13}+F_{21}+F_{23}+F_{31}+F_{32}}{F_{11}+F_{12}+F_{13}+F_{21}+F_{22}+F_{23}+F_{31}+F_{32}+F_{33}} \quad (3.11)$$

Semua hasil klasifikasi berusaha membuat model dengan nilai akurasi besar dan laju *error* kecil.

BAB IV
METODOLOGI PENELITIAN

4.1 Data Penelitian

Pada penelitian ini data yang digunakan adalah data sekunder. Data sekunder sendiri merupakan sumber data yang didapatkan peneliti dengan media perantara atau tidak secara langsung. Data sekunder yang diperoleh dari UII Career Center yaitu data *tracer study* alumni UII pada tahun 2015 di UII. Data yang diambil sebanyak 1.358 data.

4.2 Variabel dan Definisi Operasional Variabel

Pada tabel 4.1 berisi tentang penjelasan masing-masing variabel yang digunakan peneliti.

Tabel 4. 1 Definisi Variabel

Variabel	Definisi Operasional Variabel	Kategori
Jenjang Studi	Pendidikan Tinggi adalah jenjang pendidikan setelah pendidikan menengah yang mencakup program diploma, program sarjana, program magister, program doktor, program profesi, program spesialis yang diselenggarakan oleh perguruan tinggi berdasarkan kebudayaan bangsa Indonesia. (Peraturan Menteri Pendidikan dan Kebudayaan, 2014)	0 : D3/Vokasi 1 : S1/Sarjana 2 : Profesi
Fakultas	Fakultas adalah “Fakultas di Lingkungan Universitas Islam Indonesia yang berfungsi mengkoordinasi pendidikan akademik, profesi dan/atau vokasi dalam 1	0 : FE 1 : FH 2 : FIAI 3 : FK

Variabel	Definisi Operasional Variabel	Kategori
	(satu) cabang ilmu keagamaan, ilmu pengetahuan, teknologi, budaya, sastra dan/atau seni yang mencakup satu atau beberapa program/jenjang pendidikan." (Peraturan Universitas Islam Indonesia, 2014)	4 : FMIPA 5 : FPSB 6 : FTSP 7 : FTI
Indeks Prestasi Kumulatif	indeks prestasi kumulatif (IPK) yaitu jumlah perkalian nilai mutu yang diperoleh untuk setiap mata kuliah yang diambil termasuk nilai mutu ujian tesis dikalikan dengan beban Satuan Kredit Semester (SKS) dari masing-masing mata kuliah dibagi dengan SKS seluruh kegiatan akademik yang telah diikuti. (Unisys)	0 : <3 1 : 3-3,5 2 : > 3,
Jenis Kelamin	Jenis kelamin adalah atribut-atribut fisiologis dan anatomis yang membedakan antara laki-laki dan perempuan. (Wade dan Tavris, 2007)	0 : Laki-laki 1 : Perempuan
Masa Studi	Waktu maksimal yang ditempuh untuk menyelesaikan studi (Peraturan Universitas Islam Indonesia, 2014)	0 : Tidak Tepat 1 : Ya Tepat
Organisasi	Organisasi merupakan suatu kerangka hubungan yang berstruktur yang menunjukkan wewenang, tanggung jawab, dan pembagian kerja untuk menjalankan suatu fungsi tertentu.	0 : Tidak Aktif 1 : Kurang Aktif 2 : Cukup Aktif 3 : Aktif 4 : Sangat Aktif

Variabel	Definisi Operasional Variabel	Kategori
	Hubungan yang berstruktur ini disebut hirarki dan konsekuensi dari hirarki ialah adanya kategori kelompok superior dengan kelompok subordinasi. (Silalahi, 2003)	
Waktu Tunggu Bekerja	Masa dimana lulusan menunggu untuk mendapat pekerjaan dari awal lulus kuliah hingga mendapat pekerjaan. (Muslim, 2009)	

4.3 Metode Analisis Data

Dalam penelitian ini, peneliti menggunakan *software* R 3.1.1 dan *Microsoft Excel* 2013. Adapun metode analisis data yang digunakan pada penelitian ini yaitu:

1. Analisis Deskriptif, digunakan untuk mendeskripsikan karakteristik dari alumni UII yang bekerja pada 2015.
2. Analisis *machine learning* dengan metode *Naïve Bayes Clasification* dan SVM, digunakan untuk mengklasifikasikan karakteristik dari alumni UII yang bekerja tersebut ke dalam kategori kurang lancar, tidak lancar dan lancar dengan tidak lancar dan lancar.

4.4 Proses Analisis Data

Langkah-langkah analisis data pada penelitian ini adalah sebagai berikut:

1. Menentukan Kategori pada klasifikasi

Deskripsi variabel untuk 2 Kelas peneliti mengacu pada (Mutu, 2016) standar berkarya setelah lulus dari UII yaitu dikategorikan menjadi 2 kelas klasifikasi, berikut ini spesifikasinya.

Tabel 4.2 Variabel Y 2 kelas

Variabel Respon (Y)	Kriteria
Tidak Lancar = 0	Memperoleh pekerjaan dalam waktu > 3 bulan setelah lulus
Lancar = 1	Memperoleh pekerjaan dalam ≤ 3 bulan setelah lulus

2. Membagi data menjadi data *training* dan data *testing* dimana menggunakan 80% data *training* dan 20% data *testing*
3. *Support Vector Machine*

Berikut ini tahapan algoritma *Support Vector Machine*:

a. Menentukan fungsi kernel yang digunakan, pada penelitian ini menggunakan tipe kernel polynomial.
b. Menentukan nilai-nilai parameter kernel, *degree* (d) dan parameter *cost* (C) untuk melakukan optimasi.
c. Memilih parameter tebaik untuk optimasi data *training* dan untuk prediksi data *testing*
d. Mencari ketepatan prediksi menggunakan data *testing*

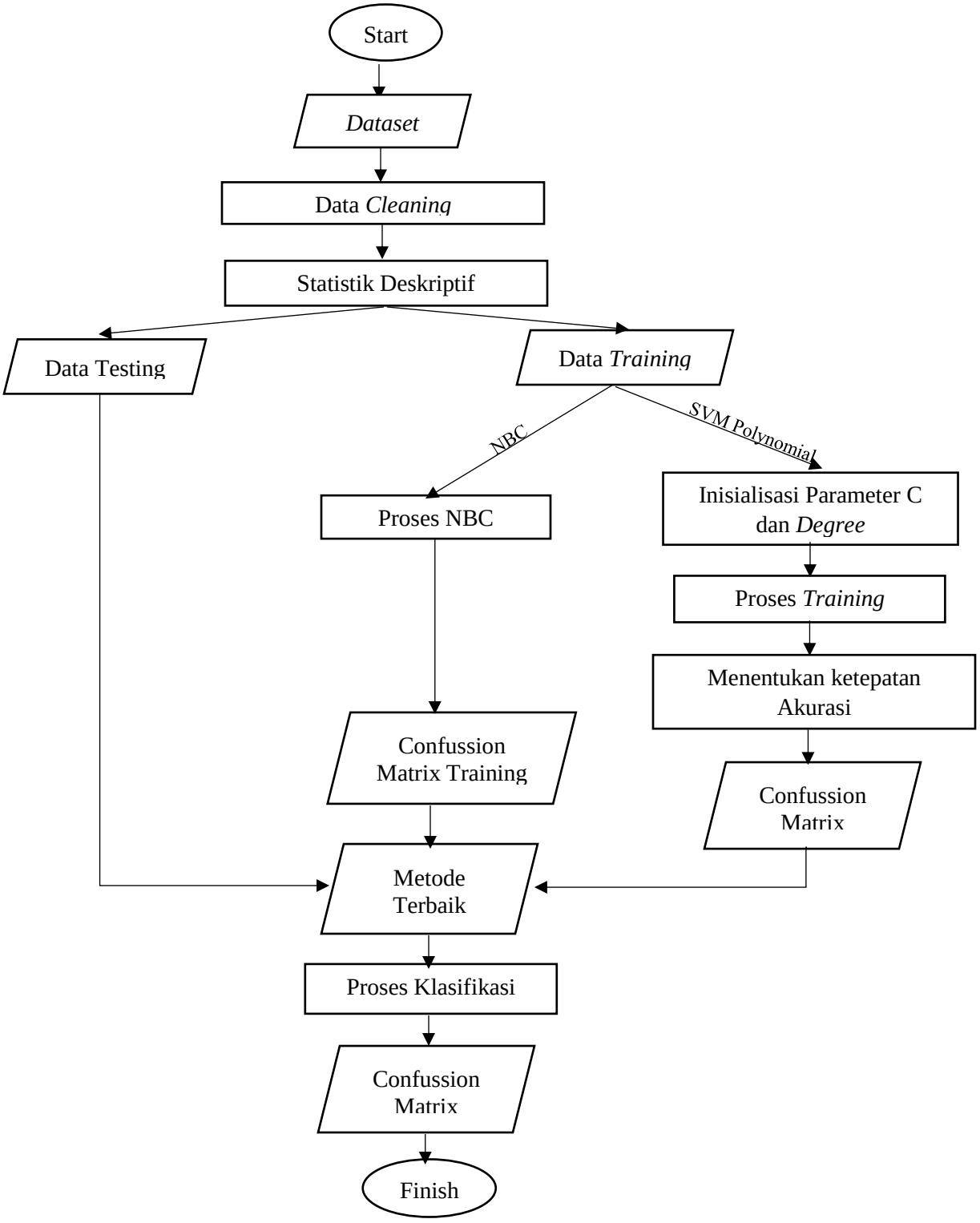
4. Naïve Bayes Classification

Berikut ini tahapan klasifikasi NBC:

a. Mencari nilai probabilitas awal (*prior probability*) peluang $P(Y)$ dan *P(X)* serta nilai $P(X_i|Y_j)$ yaitu probabilitas X_i bersyarat Y_j menggunakan data *training*.
b. Mencari hasil klasifikasi *Naïve Bayes Classification* dengan melihat hasil nilai *P(Yj|X)* paling tinggi dari setiap variabel *Y* dan menggunakan data *testing*.

4.5 Tahap Penelitian

Berikut ini tahapan penelitian menggunakan metode *Support Vector Machine* dan *Naïve Bayes Classification*.



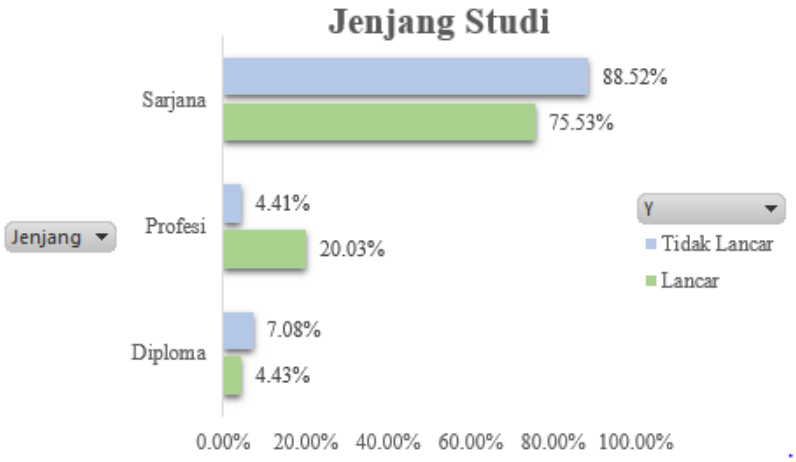
Gambar 4.1 *Diagram Alir tahap penelitian*

BAB V

HASIL DAN PEMBAHASAN

5.1 Analisis Deskriptif

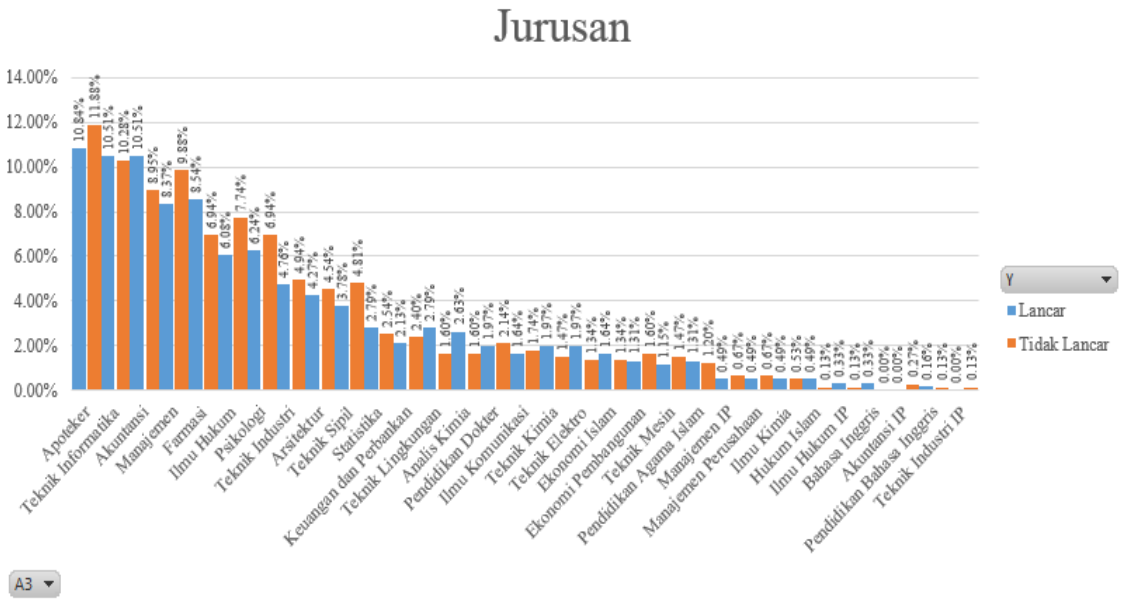
Berikut ini gambaran terkait alumni UII berdasarkan waktu tunggu dalam memperoleh pekerjaan pertama dengan data sebanyak 1.358 alumni UII.



Gambar 5.1 Presentase Jenjang Studi

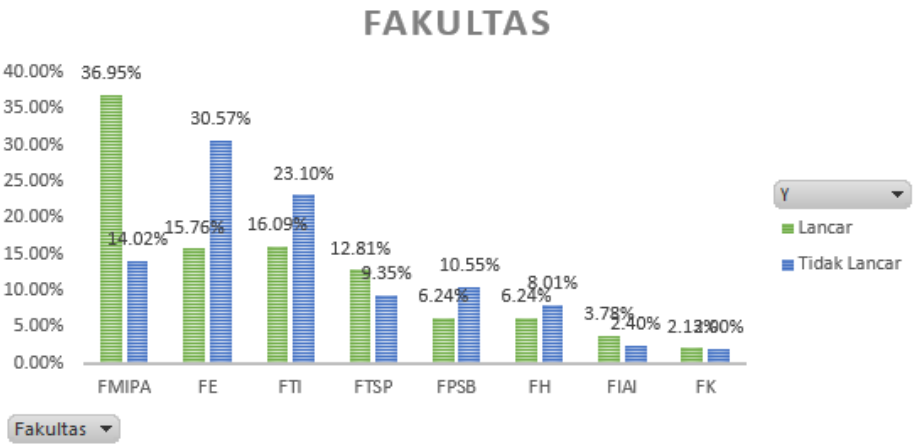
Jenjang studi pada penelitian ini dikategorikan menjadi 3 bagian yaitu Profesi, Sarjana (S1), dan D3. Berdasarkan ketiga jenjang studi tersebut jenjang studi terbanyak adalah yang berasal dari jenjang studi Sarjana (S1). Berdasarkan gambar 5.1 waktu tunggu kerja alumni tidak lancar dari jenjang studi sarjana sebesar 88,52%, jenjang studi profesi sebesar 4,41%, dan jenjang studi diploma sebesar 7.08%. Sedangkan waktu tunggu kerja alumni lancar dari jenjang studi sarjana sebesar 75,53%, jenjang studi profesi sebesar 20,03%, dan jenjang studi diploma 4,43%. Jenjang studi Sarjana ini merupakan angka yang paling tinggi dibanding jenjang studi lainnya menurut (Endarwati, 2015) Pada Agustus 2015, penduduk yang bekerja dengan lulusan sarjana hanya sedikit yaitu 8,33 %. Hal ini membuktikan bahwa lulusan sarjana kebanyakan dari mereka tidak mendapat kesempatan kerja sebab lulusan Perguruan Tinggi yang

meluluskan setiap tahunnya sehingga lapangan kerja yang terbatas tidak mampu menyerap tenaga kerja yang setiap tahunnya bertambah.



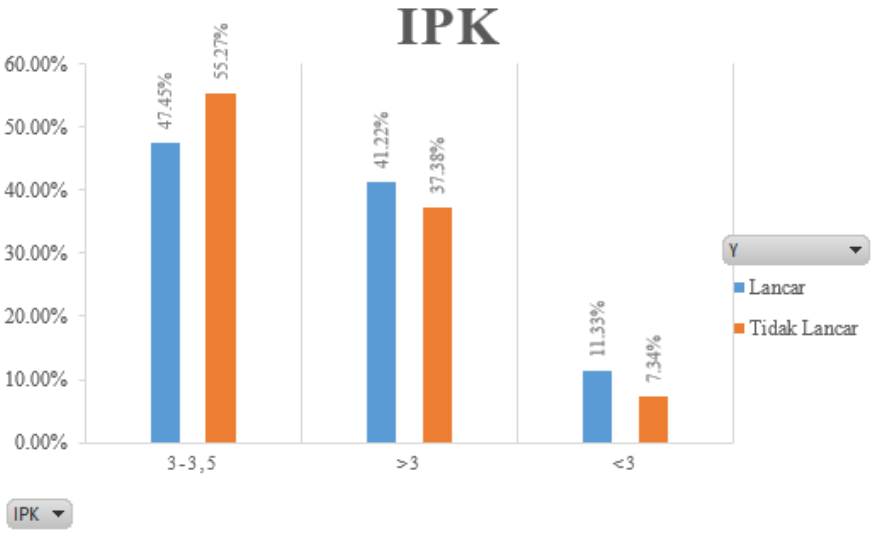
Gambar 5.2 Presentase Jurusan

Berdasarkan Jurusannya untuk waktu tunggu lancar 5 jurusan tertinggi adalah jurusan Apoteker yaitu 10,84%, Teknik Informatika yaitu 10,51%, Akuntansi yaitu 10,51%, Manajemen yaitu 8,37% dan Farmasi yaitu 8,54% sisanya untuk jurusan lainnya. Sedangkan untuk waktu tunggu tidak lancar 5 jurusan tertinggi adalah jurusan Apoteker yaitu 11,88%, Teknik Informatika yaitu 10,28%, Akuntansi yaitu 8,95%, Manajemen yaitu 9,88% dan Ilmu Hukum yaitu 7,74% sisanya untuk jurusan lainnya.



Gambar 5.3 Presentase Fakultas

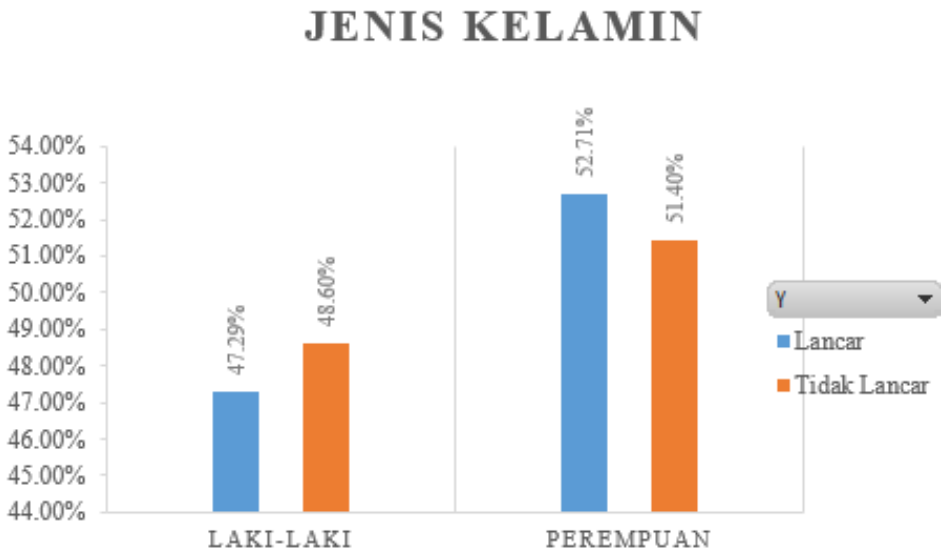
Selain jenjang studi terdapat gambaran dari fakultas yang dikategorikan menjadi 8 Fakultas berdasarkan waktu tunggu studinya Lancar yaitu dari 1.358 alumni UII sebanyak 36,95% adalah Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA), 3,78% adalah Fakultas Ilmu Agama Islam (FIAI), 12,81% adalah Fakultas Teknik Sipil dan Perencanaan (FTSP), 6,24% adalah Fakultas Psikologi dan Sosial Budaya (FPSB), 16,09% adalah Fakultas Teknik Industri (FTI), 2,13% adalah Fakultas Kedokteran (FK), 6,24% adalah Fakultas Hukum (FH), dan 15,76% adalah Fakultas Ekonomi (FE). Sedangkan untuk waktu tunggu tidak lancar berdasarkan fakultasnya yaitu FE 30,57%, FTI 23,10%, FMIPA 14,02%, FPSB 10,55%, FTSP 9,35%, FH 8,10%, FIAI 2,40%, dan FK 2,00%. Waktu tunggu kerja lancar paling tinggi berdasarkan fakultasnya yaitu FMIPA menurut (Allwar, 2017) selaku Dekan Fakultas MIPA Prodi Kimia melakukan evaluasi visi misi. Kemudian diputuskan Prodi Kimia tidak hanya menghasilkan sarjana yang siap bekerja, tetapi sarjana yang bisa melanjutkan studi di jenjang S2 hingga S3.



Gambar 5.4 Presentase IPK

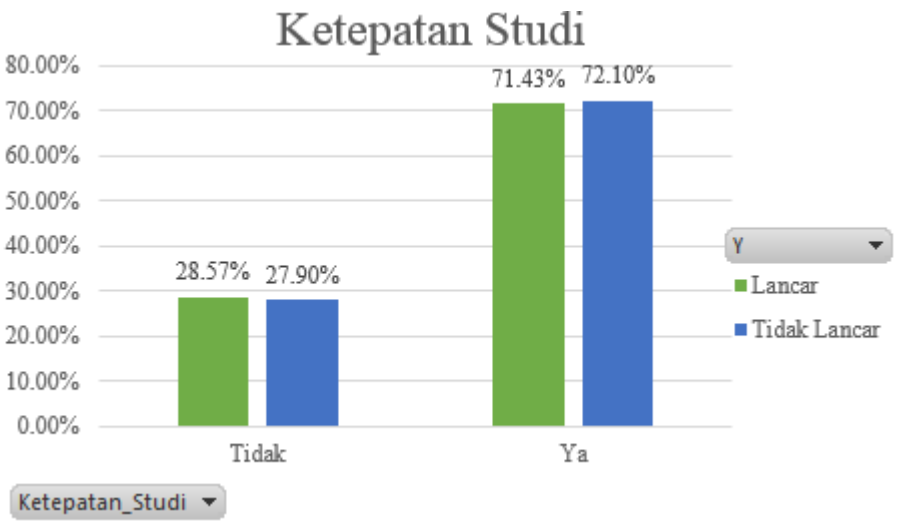
Indeks Prestasi Kumulatif (IPK) yang dikategorikan menjadi 3 bagian. Waktu tunggu kerja alumni yang lancar dengan IPK <3,00 terdapat sebanyak 11,33% sedangkan IPK 3,00-3,50 terdapat sebanyak 47,45% dan IPK >3,50 terdapat sebanyak 41,22%. Berdasarkan gambar 5.3 waktu tunggu kerja yang paling tinggi yaitu tidak

lancar dengan nilai IPK 3-3,5. Menurut (Mahaldi, 2017) Perusahaan swasta maupun BUMN dalam seleksi rekrutmen dibaca dari nilai IPK. Sehingga peluang lebih banyak nilai IPK diatas 3,0. akan menjadi pembeda diantara ribuan berkas lamaran lainnya.



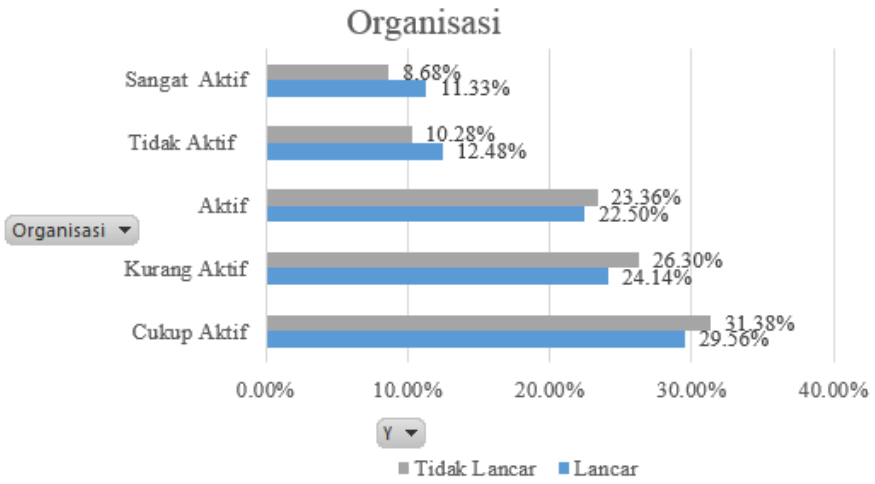
Gambar 5.5 Presentase Jenis Kelamin

Jenis kelamin yang dikategorikan menjadi 2 bagian yaitu terdapat Laki-laki dan Perempuan dimana pada gambar 5.4 menggambarkan bahwa dari 1.358 alumni dengan waktu tunggu lancar memiliki jenis kelamin perempuan adalah 52,71%. Sedangkan alumni dengan waktu tunggu lancar memiliki jenis kelamin Laki-laki adalah 47,29%. Dimana Perempuan lebih unggul 5,42% dari Laki-laki. Sedangkan pada waktu tunggu kerja untuk jenis kelamin laki-laki lebih banyak tidak lancar sebesar 48,60%. Menurut (Riyandi, 2015) data Badan Pusat Statistik (BPS) 7,3 juta orang Indonesia menganggur yang kebanyakan berjenis kelamin laki-laki bahkan pengangguran di Indonesia didominasi kalangan usia produktif dari 15 hingga 24 tahun. Penyebabnya beragam, mulai dari putus sekolah, tak melanjutkan ke jenjang perguruan tinggi, sampai sarjana yang belum mendapat pekerjaan.



Gambar 5.6 Presentase Ketepatan Studi

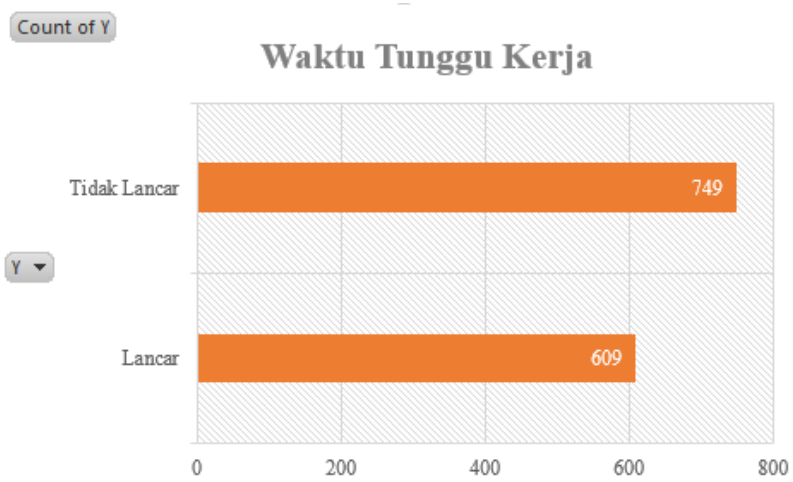
Ketepatan studi ini merupakan dalam jangka waktu yang ditetapkan bahwa alumni UII lulus dengan tepat pada waktunya maka dikatakan Ya namun bila lebih dari batas waktu yang ditentukan tersebut maka dikatakan tidak berdasarkan kategori tersebut. Pada gambar 5.5 menggambarkan bahwa alumni UII yang waktu tunggu kerjanya lancar sebesar 71,43% menjalankan studi dengan tepat waktu namun alumni UII yang waktu tunggu kerjanya tidak lancar sebesar 72,10% menjalankan studi tepat pada waktunya. Artinya lebih banyak waktu tunggu tidak lancar ini diasumsikan menurut (Amiranti, 2016) belum memiliki skill yang dibutuhkan perusahaan merupakan salah satu penyebab lulusan perguruan tinggi di Indonesia sulit mendapat pekerjaan dikutip dari Consultant Director Willis Tower Watson Indonesia, Lilis Hakim. Sedang menurut (Pusparini, 2018) ialah satu penyebab banyak alumni UII menjalankan studi dengan tepat waktu yaitu adanya workshop untuk memotivasi dan memunculkan kesadaran kepada mahasiswa dalam mengerjakan tugas akhir supaya dapat diselesaikan dengan tepat pada waktunya.



Gambar 5.7 Presentase Organisasi

Dari 1.358 alumni UII yang bekerja dengan waktu tunggu lancar sebesar 11,33% sangat aktif dalam berorganisasi, alumni UII dengan waktu tunggu lancar sebesar 22,50% aktif dalam berorganisasi, alumni UII dengan waktu tunggu lancar sebesar 29,56% cukup aktif dalam berorganisasi, alumni UII dengan waktu tunggu lancar sebesar 24,14% kurang aktif dalam berorganisasi dan alumni UII dengan waktu tunggu lancar sebesar 12,48% orang tidak aktif dalam berorganisasi. Sedangkan dengan waktu tunggu tidak lancar 8,68% sangat aktif berorganisasi, 10,28% tidak aktif berorganisasi, 23,36% aktif berorganisasi, 26,30% kurang aktif berorganisasi dan 31,38% cukup aktif berorganisasi. Menurut (Mahfud, 2014) sebagian besar interviewer lebih tertarik dengan mahasiswa yang cukup aktif di organisasi kampus dibandingkan dengan mahasiswa yang tidak aktif sama sekali di organisasi kampus. Tidak hanya itu, sangat diperlukan *soft skill* yang baik untuk lolos tahap interview yaitu seperti *skill communication*, *discussion*, *presentation*, *negotiation*, *leadership*, dan *soft skill* lainnya. *Soft-soft skill* tersebut biasanya tidak ditemukan di mata kuliah yang hanya sekedar teori.

Kemudian terdapat Waktu Tunggu Kerja yang dikategorikan menjadi 2 bagian yaitu terdapat waktu tunggu kerja Lancar dan waktu tunggu kerja tidak Lancar, berikut ini gambarannya.



Gambar 5.8 Waktu Tunggu Kerja 2 Kelas

Sedangkan untuk waktu tunggu kerja dibagi menjadi 2 bagian yaitu terdapat waktu tunggu kerja Lancar dan waktu tunggu kerja tidak Lancar. Dari 1.358 alumni UII 749 alumni UII memperoleh kerja dengan waktu tunggu tidak Lancar dan 609 alumni UII ini memperoleh kerja dengan waktu tunggu lancar.

Tabel 5.1 Tabel Analisis Deskriptif

IPK	Min.	2,290
	1st Qu.	3,180
	Median	3,420
	Mean	3,385
	3rd Qu.	3,600
	Max.	3,980
Waktu Tunggu Kerja	Min.	0,000
	1st Qu.	1,000
	Median	3,000
	Mean	4,110
	3rd Qu.	6,000
	Max.	48,000

Berdasarkan jenjang studi S1 yaitu sebesar 1.123 alumni UII. Pada Indeks Prestasi kumulatif paling terendah IPKnya sebesar 2,29, kuartil pertama sebesar 3,18, median atau nilai tengah sebesar 3,420, rata-rata IPK alumni UII sebesar 3,385, kuartil ketiga sebesar 3,60, dan IPK tertinggi sebesar 3,98. Waktu tunggu kerja paling terendah sebesar 0,000 bulan, kuartil pertama sebesar 1,000 bulan, median atau nilai tengah

sebesar 3,000 bulan, rata-rata waktu tunggu kerja alumni UII sebesar 4,110 bulan, kuartil ketiga sebesar 6,000 bulan, dan waktu tunggu kerja tertinggi sebesar 48,000 bulan.

Setelah membahas analisis deskriptif peneliti melakukan analisis menggunakan 2 kelas peneliti melakukan klasifikasi tingkat kelancaran waktu tunggu kerja alumni UII, peneliti mengacu pada (Mutu, 2016) standar berkarya setelah lulus dari UII yaitu dikategorikan menjadi 2 kelas klasifikasi, berikut ini penjelasannya.

Pada analisis klasifikasi 2 Kelas ini mengklasifikasikan data kedalam 2 kelas waktu tunggu kerja alumni UII yang berbeda yaitu ada tidak lancar dan lancar pengklasifikasian 2 kelas ini mengacu pada standart lulusan UII berkarya dalam jangka waktu tiga bulan.

5.2 Klasifikasi menggunakan *Support Vector Machine (SVM)*

Klasifikasi menggunakan metode SVM dan menggunakan fungsi Polynomial. Klasifikasi dengan metode SVM ini memerlukan ketepatan klasifikasi terbaik sehingga memerlukan nilai C (*cost*) yang sesuai untuk mengetahui hasil yang terbaik. Sebelumnya data dibagi terlebih dahulu menjadi data *training* dan data *testing*. Data yang kita gunakan tersebut dipilih secara acak dengan teknik stratified random untuk mencapai tahap optimasi supaya mendapatkan performa terbaik maka data yang digunakan untuk data *testing* berbeda dengan *training*. Pada parameter tertentu Beberapa algoritma klasifikasi terkadang dibutuhkan pengaturan.

Pengaturan fungsi kernel dibutuhkan untuk algoritma klasifikasi pada SVM. Pada penelitian ini metode Polynomial yang digunakan dalam SVM. Polynomial menghendaki, bernilai lebih dari 0. *Degree* (d) yang digunakan adalah 1, 2, 3,..., 5 Nilai Cost (c) 0,1; 1; 10; dan 100. Peneliti menggunakan metode 10-*flod cross validation* pada tahap optimasi parameter SVM karena 10-*flod cross validation* merupakan rekomendasi pemilihan model yang terbaik dan memberikan estimasi akurasi yang kurang bias dibandingkan yang lainnya. Nilai C dan *degree* ini digunakan pada data

training untuk memperoleh hasil akurasinya pada model klasifikasi. Nilai parameter C dipilih yang terbaik untuk digunakan sebagai model fungsi pemisah SVM menggunakan fungsi.

Tabel 5.2 Akurasi *Training* SVM

C	**Polynomial**				
	1	2	3	4	5
0,1	0,5994475	0,6160221	0,6233886	0,6243094	0,6233886
1	0,5994475	0,6316759	0,6298343	0,65745	0,65745
10	0,5994475	0,6325967	0,6362799	0,6593002	0,674954
100	0,5994475	0,6399632	0,6381215	0,6933702	0,6924494

Berdasarkan tabel pada data *training* ini nilai akurasi untuk c=100 dan *degree*=4 merupakan akurasi yang paling tinggi yaitu sebesar 0,693 atau 69,3%. Untuk mengukur sejauh mana kelasifier berhasil melakukan klasifikasi dengan benar merupakan fungsi dari *Testing set*. Data *testing* tidak boleh terdapat pada data *training* sehingga dapat diketahui apakah model sudah bisa dalam melakukan klasifikasi.

Tabel 5.3 Prediksi *Training* SVM

Prediction	**Reference**	
	Lancar	Tidak Lancar
Lancar	96	27
Tidak Lancar	408	555

Berdasarkan hasil prediksi pada tabel 5.15 Waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu Lancar dan waktu tunggu Lancar sebesar 96 orang diprediksikan tepat. Sedangkan waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu Lancar dan waktu tunggu tidak Lancar sebesar 27 orang diprediksikan tidak tepat. Waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu tidak Lancar dan waktu tunggu Lancar sebesar 408 orang diprediksikan tidak tepat. Terakhir yaitu waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu tidak Lancar dan waktu tunggu tidak Lancar sebesar 555 orang diprediksikan tepat Dari hasil prediksi diperoleh hasil akurasinya sebagai berikut.

Tabel 5.4 Akurasi *Training* SVM

Accuracy
0,5994475

Berdasarkan pada tabel perolehan nilai akurasi digunakan perhitungan sehingga diperoleh hasil pada tabel 5.16. Berdasarkan hasil akurasi yang dipeoleh yaitu sebesar 59,9% dan eror klasifikasi diperoleh sebesar 40,0%. Pada penelitian ini pengukuran performa menggunakan data alumni tahun 2015 dengan 1.358 *record* data. Berikut ini perhitungannya

$$Accuracy = \frac{jumlah\ prediksi\ yang\ benar}{jumlah\ prediksi\ keseluruhan}$$

$$Accuracy = \frac{651}{96+27+408+555}$$

$$Accuracy = 0{,}5994475$$

Sedangkan *error rate* atau tingkat kesalahan pada waktu tunggu kerja alumni adalah sebesar 0,4005525. Berikut ini perhitungannya.

$$Error\ Rate = \frac{jumlah\ prediksi\ yang\ salah}{jumlah\ prediksi\ keseluruhan}$$

$$Error\ Rate = \frac{435}{96+27+408+555}$$

$$Error\ Rate = 0{,}4005525$$

Setelah diperoleh hasil akurasi dari data *training*, peneliti melanjutkan mencari nilai c(*cost*) dan d(*degree*) yang paling terbesar untuk dapat memprediksikan dengan baik. Berikut ini hasil untuk polynomial dengan masing-masing nilai c(*cost*) yaitu 0.1, 1, 10, dan 100 untuk d(*degree*) yaitu 1, 2, 3, 4, dan 5.

Tabel 5.5 Akurasi *Testing* SVM

C	Polynomial				
	1	2	3	4	5
0,1	0,6875	0,6727941	0,6801471	0,6801471	0,6764706
1	0,6875	0,6654412	0,6507353	0,6433824	0,6433824
10	0,6875	0,6139706	0,6507353	0,6066176	0,625
100	0,6875	0,6139706	0,6397059	0,5661765	0,6213235

Berdasarkan tabel 5.17 pada data *testing* ini nilai akurasi untuk c=0,1 dan *degree*=1 merupakan akurasi yang paling tinggi yaitu sebesar 0,687 atau 68,7%. Untuk mengukur sejauh mana kelasifier berhasil melakukan klasifikasi dengan benar merupakan fungsi dari *Testing set*. Data *testing* tidak boleh terdapat pada data *training* sehingga dapat diketahui apakah model sudah bisa dalam melakukan klasifikasi.

Tabel 5.6 Prediksi *Testing* SVM

Prediction	**Reference**	
	Lancar	Tidak Lancar
Lancar	26	6
Tidak Lancar	79	161

Berdasarkan hasil prediksi data *testing* pada tabel 5.18 Waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu Lancar dan waktu tunggu Lancar sebesar 26 orang diprediksikan tepat. Sedangkan waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu Lancar dan waktu tunggu tidak Lancar sebesar 6 orang diprediksikan tidak tepat. Waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu tidak Lancar dan waktu tunggu Lancar sebesar 79 orang diprediksikan tidak tepat. Terakhir yaitu waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu tidak Lancar dan waktu tunggu tidak Lancar sebesar 161 orang diprediksikan tepat Dari hasil prediksi diperoleh hasil akurasinya sebagai berikut.

Tabel 5.7 Akurasi *Testing* SVM

Accuracy	**Kappa**
0,6875	0,2430592

Akurasi yang diperoleh tersebut menunjukkan bahwa dapat mengklasifikasikan waktu tunggu kerja alumni UII secara akurat dan benar dengan kelas aslinya sebesar 0,6875 atau 68,75%. Berikut ini perhitungannya

$$Accuracy = \frac{jumlah\ prediksi\ yang\ benar}{jumlah\ prediksi\ keseluruhan}$$

$$Accuracy = \frac{187}{26+6+79+161}$$

$$Accuracy = 0{,}6875$$

Sedangkan *error rate* atau tingkat kesalahan pada waktu tunggu kerja alumni adalah sebesar 0,3125 berikut ini perhitungannya.

$$Error\ Rate = \frac{jumlah\ prediksi\ yang\ salah}{jumlah\ prediksi\ keseluruhan}$$

$$Error\ Rate = \frac{85}{26+6+79+161}$$

$$Error\ Rate = 0{,}3125$$

5.2.1. Perhitungan *Support Vector Machine*

– Menghitung K(u,z) : dengan u=(1,1) dan z=(1,-1)

$k(u = (1,1), z = (1,-1)) = (((u_1.z_1) + (u_2.z_2)) + 1)^2$

$= ((u_1.z_1) + (u_2.z_2))2 + 2((u_1.z_1) + (u_2.z_2)).1 + 1^2$

$= (u_1.z_1)2 + 2(u_1.z_1)(u_2.z_2) + (u_2.z_2)2 + 2(u_1.z_1) + 2(u_2.z_2) + 1$

$$= \begin{pmatrix} {u_1}^2 \\ \sqrt{2}u_1u_2 \\ {u_2}^2 \\ \sqrt{2}u_1 \\ \sqrt{2}u_2 \\ 1 \end{pmatrix} \begin{pmatrix} {z_1}^2 \\ \sqrt{2}z_1z_2 \\ {z_2}^2 \\ \sqrt{2}z_1 \\ \sqrt{2}z_2 \\ 1 \end{pmatrix}$$

$= \phi(u).\phi(z)$

$= {u_1}^2{z_1}^2 + 2u_1u_2z_1z_2 + (u_2z_2)^2 + 2(u_1.z_1) + 2(u_2z_2) + 1$

$= (u_1z_1)^2 + 2(u_1z_1)(u_2z_2) + (u_2z_2)^2 + 2(u_1z_1) + 2(u_2z_2) + 1$

– K(u,z) : dengan u=(1,1) dan z=(1,-1)

$k(U=(1,1),Z=(1,-1)) = (((1.1)+(1.(-1)))+1)^2$

$= ((1.1)+(1.(-1)))^2+2((1.1)+(1.(-1))).1 + 1^2$

$= (1.1)^2 + 2(1.1)(1.(-1)) + (1.(-1))^2 + 2(1.1) + 2(1.(-1)) + 1$

$= 1 - 2 + 1 + 2 - 2 + 1 = 1$

$$= \begin{pmatrix} {u_1}^2 \\ \sqrt{2}u_1u_2 \\ {u_2}^2 \\ \sqrt{2}u_1 \\ \sqrt{2}u_2 \\ 1 \end{pmatrix} \begin{pmatrix} {z_1}^2 \\ \sqrt{2}z_1z_2 \\ {z_2}^2 \\ \sqrt{2}z_1 \\ \sqrt{2}z_2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1^2 \\ \sqrt{2}.1.1 \\ 1^2 \\ \sqrt{2}.1 \\ \sqrt{2}.1 \\ 1 \end{pmatrix} \begin{pmatrix} 1^2 \\ \sqrt{2}.1.(-1) \\ 1^2 \\ \sqrt{2}.1 \\ \sqrt{2}.(-1) \\ 1 \end{pmatrix}$$

= 1-2+1+2-2+1=1

$= \phi(u).\phi(z)$

$= u_1^2 z_1^2 + 2u_1u_2z_1z_2 + (u_2z_2)^2 + 2(u_1.z_1) + 2(u_2z_2) + 1$

$= (u_1z_1)^2 + 2(u_1z_1)(u_2z_2) + (u_2z_2)^2 + 2(u_1z_1) + 2(u_2z_2) + 1$

– Menghitung matrik kernel K(x,xi) = ϕ(x).ϕ(xi)

Tabel 5.8 Menghitung Matrik kernel

x_1	x_1	$K(1,0) = (x_1.x_1 + 1)^2 = (1.1 + 1.(0) +1)^2 = 2^2 = 4$
	x_2	$K(1,1) = (x_1.x_2 + 1)^2 = (1.1 + 1.1 +1)^2 = 3^2 = 9$
	x_3	$K(1,1) = (x_1.x_3 + 1)^2 = (1.1 + 1.1 +1)^2 = 3^2 = 9$
	x_4	$K(1,1) = (x_1.x_4 + 1)^2 = (1.1 + 1.1 +1)^2 = 3^2 = 9$
x_2	x_1	$K(0,0) = (x_2.x_1 + 1)^2 = ((0).(0) + (0).(0) +1)^2 = 1^2 = 1$
	x_2	$K(0,1) = (x_2.x_2 + 1)^2 = ((0).(0) + (0).1 +1)^2 = 1^2 = 1$
	x_3	$K(0,1) = (x_2.x_3 + 1)^2 = ((0).(0) + (0).1 +1)^2 = 1^2 = 1$
	x_4	$K(0,1) = (x_2.x_4 + 1)^2 = ((0).(0) + (0).1 +1)^2 = 1^2 = 1$
X_3	x_1	$K(1,0) = (x_3.x_1 + 1)^2 = (1.1 + 1.(0) +1)^2 = 2^2 = 4$
	x_2	$K(1,1) = (x_3.x_2 + 1)^2 = (1.1 + 1.1 +1)^2 = 3^2 = 9$
	x_3	$K(1,1) = (x_3.x_3 + 1)^2 = (1.1 + 1.1 +1)^2 = 3^2 = 9$
	x_4	$K(1,1) = (x_3.x_4 + 1)^2 = (1.1 + 1.1 +1)^2 = 3^2 = 9$
x_4	x_1	$K(0,0) = (x_4.x_1 + 1)^2 = ((0).(0) + (0).(0) +1)^2 = 1^2 = 1$
	x_2	$K(0,1) = (x_4.x_2 + 1)^2 = ((0).(0) + (0).1 +1)^2 = 1^2 = 1$
	x_3	$K(0,1) = (x_4.x_3 + 1)^2 = ((0).(0) + (0).1 +1)^2 = 1^2 = 1$
	X_4	$K(0,1) = (x_4.x_4 + 1)^2 = ((0).(0) + (0).1 +1)^2 = 1^2 = 1$

– Didapatkan matrik kernel K dengan ukuran (NxN)

K(x,xi) =

4	9	9	9
1	1	1	1
4	9	9	9
1	1	1	1

– Memaksimalkan Ld

$$Ld = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i.j} \alpha_i \alpha_j y_i y_j (x_i . x_j)$$

$$\text{Ld} = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4) - \frac{1}{2}\begin{pmatrix} 4\alpha_1\alpha_1 + 9\alpha_1\alpha_2 + 9\alpha_1\alpha_3 - 9\alpha_1\alpha_4 + \\ 1\alpha_2\alpha_1 + 1\alpha_2\alpha_2 + 1\alpha_2\alpha_3 - 1\alpha_2\alpha_4 + \\ 4\alpha_3\alpha_1 + 9\alpha_3\alpha_2 + 9\alpha_3\alpha_3 - 9\alpha_3\alpha_4 + \\ 1\alpha_4\alpha_1 + 1\alpha_4\alpha_2 + 1\alpha_4\alpha_3 - 1\alpha_4\alpha_4 \end{pmatrix}$$

$$\text{Ld} = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4) - \frac{1}{2}\begin{pmatrix} 4{\alpha_1}^2 + 10\alpha_1\alpha_2 + 13\alpha_1\alpha_3 - 8\alpha_1\alpha_4 + \\ 1\alpha_2\alpha_2 + 10\alpha_2\alpha_3 + \\ 9{\alpha_3}^2 - 8\alpha_3\alpha_4 - 1{\alpha_4}^2 \end{pmatrix}$$

Syarat 1 : $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \geq 0$

Syarat 2 : $\alpha_1+\alpha_2+\alpha_3-\alpha_4 = 0$

Nilai max Ld $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$ = 0,125 sehingga nilai Ld = 0,25

– Hitung nilai w dan b

$$w = \sum_{i=1}^{N} \alpha_i y_i \phi(X_i)$$

$$w = \sum_{i=1}^{4} \alpha_i y_i \phi(X_i) = \alpha_1 y_1 \phi(X_1) + \alpha_2 y_2 \phi(X_2) + \alpha_3 y_3 \phi(X_3) + \alpha_4 y_4 \phi(X_4)$$

$$w = 0{,}125\begin{pmatrix} {x_1}^2 = 1^2 \\ \sqrt{2}x_1x_2 = \sqrt{2}.1.0 \\ {x_2}^2 = 0^2 \\ \sqrt{2}x_1 = \sqrt{2}.1 \\ \sqrt{2}x_2 = \sqrt{2}.0 \\ 1 \end{pmatrix} + 0{,}125\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + 0{,}125\begin{pmatrix} 1 \\ \sqrt{2} \\ 0 \\ \sqrt{2} \\ 0 \\ 1 \end{pmatrix} - 0{,}125\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0{,}25 \\ 0{,}177 \\ 0 \\ 0{,}354 \\ 0 \\ 0{,}25 \end{pmatrix}$$

$$w = \sum_{i=1}^{N} \alpha_i y_i \phi(X_i) = \begin{pmatrix} 0{,}25 \\ 0{,}177 \\ 0 \\ 0{,}354 \\ 0 \\ 0{,}25 \end{pmatrix}$$

– Pilih support vector dari kelas +1 dan -1 untuk menghitung b

$$b = -\frac{1}{2}(w.x^{+} + w.x^{-}) = -\frac{1}{2}\begin{pmatrix}0{,}25\\0{,}177\\0\\0{,}354\\0\\0{,}25\end{pmatrix}\begin{pmatrix}1\\\sqrt{2}\\0\\\sqrt{2}\\0\\1\end{pmatrix} + \begin{pmatrix}0{,}25\\0{,}177\\0\\0{,}354\\0\\0{,}25\end{pmatrix}\begin{pmatrix}0\\0\\0\\0\\0\\1\end{pmatrix}$$

$$= \Big((0{,}25)(1) + (0{,}177)\big(\sqrt{2}\big) + (0{,}354)\big(\sqrt{2}\big) + (0{,}25)(1)\Big) + 0{,}25 = 1{,}5$$

– Maka model SVM siap digunakan untuk proses klasifikasi.

$$f\big(\phi(x)\big) = sign\big(w.\phi(x) + b\big) = sign\left(\sum_{i=1}^{N} \alpha_i y_i \phi(x_i).\phi(x) + b\right)$$

– Misalkan data uji/ data test $x_t = (0,0)$ maka $K(x_i,x_t) = \phi(x_i).\ \phi(x_t)$

Tabel 5.9 Menghitung dengan data uji

x_t=(0,0)	x_1	$K(1,t) = (x_1.x_t + 1)^2 = (0.1 + 0.0 + 1)^2 = 1^2 = 1$	(0,125)(1)
	x_2	$K(2,t) = (x_2.x_t + 1)^2 = (0.0 + 0.1) + 1)^2 = 1^2 = 1$	(0,125)(1)
	x_3	$K(3,t) = (x_3.x_t + 1)^2 = (0.1 + 0.1 + 1)^2 = 1^2 = 1$	(0,125)(1)
	x_4	$K(4,t) = (x_4.x_t + 1)^2 = (0.0 + 0.1 + 1)^2 = 1^2 = 1$	(-0,125)(1)

– Hasil *hyperplane* dari model :

$$f\big(\Phi(x)\big) = \text{sign}(w.\Phi(x_t) + b) = \text{sign}(\textstyle\sum_{i=1}^{ns} \alpha_i y_i \Phi(x_i).\Phi(x_t) + b)$$

$$= \text{sign}(\textstyle\sum_{i=1}^{ns} \alpha_i y_i K(x_i, x_t) + b)$$

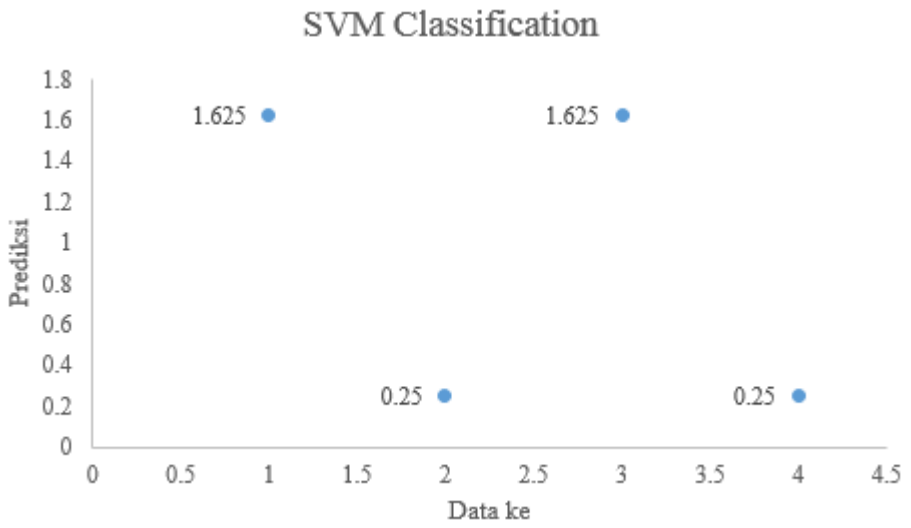
$$= \text{sign}(\textstyle\sum_{i=1}^{ns} \alpha_i y_i K(0,0) + b)$$

$$= \text{sign}(0{,}125 + 0{,}125 + 0{,}125 - 0{,}125 + 0)$$

$$= \text{sign}(0{,}25) \rightarrow 1$$

Jadi data $x_t = (0,0)$ tersebut masuk ke kelas positif

Berdasarkan dari contoh tersebut ilustrasi gambar *hyperplane* dengan menggunakan 4 data. Variabel yang digunakan jenis kelamin dan ketepatan studi dengan klasifikasi waktu tunggu kerja lancar dan tidak lancar. maka digambarkan dalam bentuk grafik sebagai berikut.



Gambar 5.9 *Hyperplane*

Berdasarkan gambar 5.9 bahwa keempat data tersebut termasuk dalam kelas positif +1 yang artinya bahwa termasuk dalam kategori lancer. *Hyperplane* tidak muncul dikarenakan tidak ada yang termasuk dalam kelas negative -1 yang artinya tidak ada pemisah kelas pada keempat data tersebut.

5.3 Klasifikasi Menggunakan *Naïve Bayes Classification*

Berdasarkan permasalahan 2 kelas dilanjutkan dengan membangun sebuah sistem untuk menentukan waktu tunggu kerja alumni UII dengan menggunakan analisis *Naïve Bayesian Classification*. Metode *Naïve Bayes Classification* ini hampir serupa dengan metode *Support Vector Machine* yaitu menggunakan data *training* terlebih dahulu kemudian menggunakan data *testing* untuk melihat nilai akurasi dari metode tersebut.

Tabel 5.10 A-priori probabilities

Lancar	Tidak Lancar
0,4485294	0,5514706

Berdasarkan data jumlah data latih diperoleh probabilitas waktu tunggu alumni UII memperoleh pekerjaan yaitu untuk waktu tunggu lancar sebesar 0,448, dan untuk

probabilitas waktu tunggu tidak lancar sebesar 0,551. Berikut ini hasil prediksi dari data *training*.Berikut ini pembahasannya.

Tabel 5.11 Prediksi *Training* NBC

Prediction	**Reference**	
	Lancar	Tidak Lancar
Lancar	131	62
Tidak Lancar	357	538

Berdasarkan hasil prediksi data *training* pada tabel 5.20 Waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu Lancar dan waktu tunggu Lancar sebesar 131 orang diprediksikan tepat. Sedangkan waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu Lancar dan waktu tunggu tidak Lancar sebesar 62 orang diprediksikan tidak tepat. Waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu tidak Lancar dan waktu tunggu Lancar sebesar 357 orang diprediksikan tidak tepat. Terakhir yaitu waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu tidak Lancar dan waktu tunggu tidak Lancar sebesar 538 orang diprediksikan tepat Dari hasil prediksi diperoleh hasil akurasinya sebagai berikut.

Tabel 5.12 Akurasi *Training* NBC

Accuracy
0,6148897

Dari hasil prediksi diperoleh nilai akurasi dari analisis *Naïve Bayesian Clasification* dengan menggunakan 2 kelas yaitu sebesar 0,6148897 atau 61,48% dengan menggunakan perhitungan dari persamaan berikut ini.

$$Accuracy = \frac{jumlah\ prediksi\ yang\ benar}{jumlah\ prediksi\ keseluruhan}$$

$$Accuracy = \frac{669}{131+62+357+538}$$

$$Accuracy = 0{,}6148897$$

Sedangkan *error rate* atau tingkat kesalahan pada waktu tunggu kerja alumni adalah sebesar 0,3851102 berikut ini perhitungannya.

$$Error\ Rate = \frac{jumlah\ prediksi\ yang\ salah}{jumlah\ prediksi\ keseluruhan}$$

$$Error\ Rate = \frac{419}{131+62+357+538}$$

$$Error\ Rate = 0{,}3851102$$

Setelah akurasi dari data *training* diketahui selanjutnya peneliti memprediksi dari data *testing* dan memperoleh nilai akurasinya.

Tabel 5.13 Prediksi *Testing* NBC

Prediction	**Reference**	
	Lancar	Tidak Lancar
Lancar	25	13
Tidak Lancar	96	136

Berdasarkan hasil prediksi data *testing* pada tabel 5.22 Waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu Lancar dan waktu tunggu Lancar sebesar 25 orang diprediksikan tepat. Sedangkan waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu Lancar dan waktu tunggu tidak Lancar sebesar 13 orang diprediksikan tidak tepat. Waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu tidak Lancar dan waktu tunggu Lancar sebesar 96 orang diprediksikan tidak tepat. Terakhir yaitu waktu tunggu alumni UII memperoleh pekerjaan dengan waktu tunggu tidak Lancar dan waktu tunggu tidak Lancar sebesar 136 orang diprediksikan tepat Dari hasil prediksi diperoleh hasil akurasinya sebagai berikut.

Tabel 5.14 Akurasi *Testing* NBC

Accuracy	**Kappa**
0,5963	0,1276

Dari hasil prediksi diperoleh nilai akurasi dari analisis *Naïve Bayesian Clasificatio*n dengan menggunakan 2 kelas yaitu sebesar 0,5963 atau 59,63%. Berdasarkan hasil *output* yang dihasilkan memiliki nilai kappa 0,1276 yang artinya kemungkinan hasil klasifikasi tersebut dapat digunakan dengan menggunakan perhitungan dari persamaan berikut ini.

$$Accuracy = \frac{jumlah\ prediksi\ yang\ benar}{jumlah\ prediksi\ keseluruhan}$$

$$Accuracy = \frac{161}{25+13+96+136}$$

$$Accuracy = 0{,}5963$$

Sedangkan *error rate* atau tingkat kesalahan pada waktu tunggu kerja alumni adalah sebesar 0,4037 berikut ini perhitungannya.

$$Error\ Rate = \frac{jumlah\ prediksi\ yang\ salah}{jumlah\ prediksi\ keseluruhan}$$

$$Error\ Rate = \frac{109}{25+13+96+136}$$

$$Error\ Rate = 0{,}4037$$

5.4 Perbandingan Analisis SVM dengan NBC

Setelah dilakukan kedua pengujian analisis klasifikasi dari SVM dan NBC maka peneliti melakukan perbandingan diantara kedua metode.

Tabel 5.15 Perbandingan Klasifikasi SVM dan NBC

SVM	Kappa	NBC	Kappa
68,75%	0,2430592	59,63%	0,1276

Berdasarkan hasil perbandingan dua metode klasifikasi yaitu SVM lebih besar dibandingkan dengan NBC diperoleh akurasi dari metode SVM yaitu sebesar 68,75% dan akurasi untuk metode NBC yaitu sebesar 59,63% sehingga diantara keduanya ialah metode SVM yang paling baik. Berdasarkan nilai kappa yang dihasilkan sebesar 0,2430592 yang artinya kemungkinan hasil klasifikasi tersebut dapat digunakan dengan nilai akurasinya yang paling tinggi adalah SVM menggunakan nilai *cost* 0,1 dan *degree* 1 dieperoleh 0,6875 atau 68,75% artinya waktu tunggu kerja mahasiswa UII pada tahun 2015 dapat diklasifikasikan dengan tingkat akurasi sebesar 68,75% menggunakan metode SVM nilai *cost* 0,1 dan *degree* 1 dapat diklasifikasikan dengan baik. Hasil akurasi yang diperoleh sebesar 68,75% karena diasumsikan ada variabel lain yang dapat mempengaruhi hasil akurasinya.

BAB VI
PENUTUP

6.1 Kesimpulan

Berdasarkan rumusan masalah yang telah dijelaskan sebelumnya, sehingga diperoleh kesimpulan untuk menjawab rumusan masalah yang ada, yaitu:

1. Gambaran dari faktor yang mendukung waktu tunggu kerja alumni UII tidak lancar dari jenjang studi sarjana sebesar 88,52%. Waktu tunggu kerja lancar paling tinggi berdasarkan fakultasnya yaitu FMIPA. Nilai IPK dengan waktu tunggu kerja tidak lancar yang paling tinggi IPK dengan nilai 3-3,5. Jenis kelamin perempuan memiliki waktu tunggu lancar sebesar 52,71%. Sebesar 72.10% menjalankan studi tepat pada waktunya dengan waktu tunggu kerjanya tidak lancar. Waktu tunggu kerja paling terendah sebesar 0,000 bulan, rata-rata waktu tunggu kerja alumni UII sebesar 4,110 bulan, dan waktu tunggu kerja tertinggi sebesar 48,000 bulan.
2. Hasil akurasi dari kedua metode yang terbaik hasil klasifikasi 2 kelas nilai akurasinya yang terbaik adalah dengan metode SVM menggunakan nilai *cost* 0,1 dan *degree* 1 sehingga diperoleh klasifikasi yang terbaik sebesar 0,6875 atau 68,75%.

6.2 Saran

1. Saran untuk penelitian selanjutnya yaitu menambahkan variabel yang lebih banyak supaya dapat dilihat hasil keakuratan dengan menggunakan klasifikasi lebih baik.
2. Saran untuk perguruan tinggi diharapkan dapat bekerjasama dengan berbagai macam perusahaan dan menjalin MOU yang pasti supaya mempermudah dalam menghasilkam lulusan banyak yang bekerja dengan waktu tunggu lancar.

DAFTAR PUSTAKA

Allwar. (2017). *Berita Kampus*. Diakses pada Prodi Kimia Hasilkan Lulusan Siap Kerja Dan Studi Lanjut: http://science.uii.ac.id/prodi-kimia-uii-hasilkan-lulusan-siap-kerja-dan-studi-lanjut/

Amiranti, C. C. (2016). *Kompas.com*. Diakses pada Jangan Asal Kuliah, Ijazah Sarjana Tak Cukup Buat Kerja!: https://edukasi.kompas.com/read/2016/11/25/06310081/jangan.asal.kuliah.ijazah.sarjana.saja.tak.cukup.buat.kerja.

Bhawiyuga, M. F. (2018). Analisis Perbandingan Akurasi Deteksi Serangan Pada Jaringan Komputer Dengan Metode Naïve Bayes Dan Support Vector Machine (SVM). *JPTIIK*.

Center, U. C. (2017). Diakses pada Tracer Study UII: https://tracer.uii.ac.id/tentang-kami.html

Darmawan, Baziedy Aditya. 2017 Pentingnya Tracer Study dalam Mengukur Kualitas Proses Pendidikan Tinggi Diakses pada 12 Mei 2018 dari https://career.uii.ac.id/detailnews/165/pentingnya-tracer-study-dalam-mengukur-kualitas-proses-pendidikan-tinggi.html

Dwihartanti, M. (2015). Diakses pada Laporan Tracer Study Alumni: http://staff.uny.ac.id/sites/default/files/penelitian/muslikhah-dwihartanti-mpd/laporan-tracer-study-alumni.pdf

Endarwati, S. (2015). *Kompasiana*. Diakses pada Lulusan Sarjana Dominasi Pengangguran di Indonesia, Siapkah Hadapi MEA?: https://www.kompasiana.com/sulih043/lulusan-sarjana-dominasi-pengangguran-di-indonesia-siapkah-hadapi-mea_56654fba93937334079cb384

Fakhriyani, I. (2016). *Perbandingan Algoritma Naive Bayes Dan Support Vector Machine Dalam Seleksi Kelulusan Pemberkasan Beasiswa BPP-PPA.* Jakarta: Fakultas Teknik Universitas Negeri Jakarta.

Faramudhita, S. (2018). *Analisis Klasifikasi Menggunakan Metode Regresi Logistik Ordinal dan Naive Bayes Pada Data Alumni UNILA Tahun 2016.* Bandar Lampung.

Fitri, A. (2010). *Hubungan Motivasi dengan Lamanya Masa Tunggu Kerja pada Lulusan Fakultas Keperawatan USU Periode 2008 dan 2009.* Medan.

Furey TS, C. N. (2000). Support Vector Machine Classification dan Validation Of Cancer Tissue Samples Using Microarray Expression Data . *Bioinformatics*.

Han J., M. K. (2000). *Data Mining: Concepts And Techniques.* New York: Morgan-Kaufman.

Hoyyi, D. I. (2016). Analisis Klasifikasi Masa Studi Mahasiswa Prodi Statistika UNDIP Dengan Metode Support Vector Machine (SVM) Dan ID3 (Iterative Dichotomiser 3). *Media Statistika*.

Okezone. (2018). Angka Pengangguran Indonesia Capai Level Terendah Sepanjang Sejarah. *Okezone*.

Mahaldi, H. (2017). *Diary Kantor*. Retrieved April 14 , 2018, from Kenapa Kita Harus Meraih IPK 3 Keatas Sebelum Lulus?: https://diarykantor.com/kenapa-kita-harus-meraih-ipk-3-keatas-sebelum-lulus/

Mahfud, A. (2014). *Mahfud Berbagi*. Diakses pada Kuliah Itu Mesti berorganisasi: http://aunurmahfud.blogspot.co.id/2014/04/kuliah-itu-mesti-berorganisasi.html

Mamiek. (2016). *Akuntansi UII Persiapkan Mahasiswa Bersertifikasi Profesi Internasional*. Diakses pada http://accounting.uii.ac.id/akuntansi-uii-persiapkan-mahasiswa-bersertifikasi-profesi-internasional/

Mashuri, R. T. (2012). Analisis Statistik Tentang Faktor-Faktor Yang Mempengaruhi Waktu Tunggu Kerja Fresh graduate Di Jurusan Statistika Institut Teknologi Sepuluh Nopember (ITS) Dengan Metode Regresi Logistik Ordinal.

Muslim, I. N. (2009). Hubungan Motivasi Bekerja Dengan Waktu Tunggu Kerja Pada Lulusan Program Studi Pendidikan Tata Rias.

Mutu, B. P. (2016). *Standar Mercy Of God*. Diakses pada http://bpm.uii.ac.id/wp-content/uploads/2016/10/Poster-Sasaran-Mutu-QA-Board_final_Fakultas.pdf

Octaviani, Pusphita Anna. Y. W. (2014). Penerapan Metode Klasifikasi Support Vector Machine (SVM) Pada Data Akreditasi Sekolah Dasar (SD) Di Kabupaten Magelang. *Jurnal Gaussian*, 811-820Prasetyo, E. (2012). *Data Mining: Konsep dan Aplikasi Menggunakan MATLAB.* Yogyakarta: C.V Andi Offset.

Pusparini, Martini Dwi. (2018). Diakses pada Public Hearing Motivasi Mahasiswa Percepat Penyelesaian Skripsi: https://www.uii.ac.id/public-hearing-motivasi-mahasiswa-percepat-penyelesaian-skripsi/

Ridwan, M. S. (2013). Penerapan Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classification Kelasifier. *Jurnal EECCIS*.

Ristekdikti. (2016). Diakses pada Pelaksanaan Tracer Study Tingkat Perguruan Tinggi : http://belmawa.ristekdikti.go.id/2016/08/02/pelaksanaan-tracer-study-tingkat-perguruan-tinggi/

Riyandi, S. (2015). Retrieved April 16, 2018, from 7,3 Juta orang Indonesia masih menganggur, kebanyakan laki-laki: https://www.merdeka.com/uang/73-juta-orang-indonesia-masih-menganggur-kebanyakan-laki-laki.html

Saleh, A. (2015). *Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga.* Medan: Universitas Potensi Utama.

Santosa, B. (2007). *Data Mining: Teknik Pemanfaatan Data Untuk Keperluan Bisnis.* Yogyakarta: Graha Ilmu.

Silalahi. (2003). *Studi Tentang Ilmu Administrasi Konsep, Teori, dan Dimensi.* Bandung: C.V Sinar Baru.

Tavris, C. W. (2007). *Psikologi.* Jakarta: Erlangga.

Unisys. (n.d.). Diakses pada Pendidikan: https://unisys.uii.ac.id/cetak.asp?u=912&b=I&v=2&j=I&id=9&owner=912

LAMPIRAN

LAMPIRAN 1. *Crosstabulation* Data

a. Variabel IPK 2 Kelas

Count of Y	Column Labels		
Row Labels	Lancar	Tidak Lancar	Grand Total
<3	69	55	124
3-3,5	289	414	703
>3	251	280	531
Grand Total	**609**	**749**	**1358**

b. Variabel Jenjang Studi 2 Kelas

Count of Y	Column Labels		
Row Labels	Lancar	Tidak Lancar	Grand Total
Sarjana	460	663	1123
Diploma	27	53	80
Profesi	122	33	155
Grand Total	**609**	**749**	**1358**

c. Variabel Fakultas 2 Kelas

Count of Y	Column Labels		
Row Labels	Lancar	Tidak Lancar	Grand Total
FTSP	78	70	148
FK	13	15	28
FTI	98	173	271
FPSB	38	79	117
FH	38	60	98
FMIPA	225	105	330
FE	96	229	325
FIAI	23	18	41
Grand Total	**609**	**749**	**1358**

d. Variabel Jenis Kelamin 2 Kelas

Count of Y	Column Labels		
Row Labels	Lancar	Tidak Lancar	Grand Total
Laki-Laki	288	364	652
Perempuan	321	385	706
Grand Total	**609**	**749**	**1358**

e. Variabel Ketepatan Studi 2 Kelas

Count of Y	Column Labels		
Row Labels	Lancar	Tidak Lancar	Grand Total
Tidak	174	209	383
Ya	435	540	975
Grand Total	**609**	**749**	**1358**

f. Variabel Organisasi 2 Kelas

Count of Y	Column Labels		
Row Labels	Lancar	Tidak Lancar	Grand Total
Aktif	137	175	312
Cukup Aktif	180	235	415
Kurang Aktif	147	197	344
Sangat Aktif	69	65	134
Tidak Aktif	76	77	153
Grand Total	**609**	**749**	**1358**

g. Waktu Tunggu Kerja 2 Kelas

Row Labels	Count of Y
Lancar	609
Tidak Lancar	749
Grand Total	**1358**

LAMPIRAN 2. Hasil Analisi Klasifikasi SVM dan NBC

a. Jumlah Data NBC 2 Kelas

```
Lancar Tidak Lancar
   609         749
> nrow(dataTrain)
[1] 1088
> nrow(dataTest)
[1] 270
```

b. Nilai Probabilitas NBC 2 Kelas

```
> prop.table(table(dataTrain$Y))

      Lancar Tidak Lancar
   0.4485294    0.5514706
> prop.table(table(dataTest$Y))

      Lancar Tidak Lancar
   0.4481481    0.5518519
Naive Bayes Classifier for Discrete Predictors

Call:
naiveBayes.default(x = X, y = Y, laplace = laplace)

A-priori probabilities:
Y
      Lancar Tidak Lancar
   0.4485294    0.5514706
```

c. Hasil Prediksi training 25 teratas NBC 2 Kelas

```
 [1] Tidak Lancar Tidak Lancar Lancar       Tidak Lancar Tidak Lancar
 [6] Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar
[11] Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar
[16] Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar
[21] Tidak Lancar Lancar       Lancar       Tidak Lancar Lancar
Levels: Lancar Tidak Lancar
```

d. Hasil Confussion Matrix NBC 2 Kelas

```
Confusion Matrix and Statistics

             TRUE
prediction    Lancar Tidak Lancar
  Lancar          25           13
  Tidak Lancar    96          136

               Accuracy : 0.5963
                 95% CI : (0.5351, 0.6553)
    No Information Rate : 0.5519
    P-Value [Acc > NIR] : 0.07929

                  Kappa : 0.1276
 Mcnemar's Test P-Value : 4.024e-15

            Sensitivity : 0.20661
            Specificity : 0.91275
         Pos Pred Value : 0.65789
         Neg Pred Value : 0.58621
             Prevalence : 0.44815
         Detection Rate : 0.09259
   Detection Prevalence : 0.14074
      Balanced Accuracy : 0.55968

       'Positive' Class : Lancar
```

e. Hasil Prediksi dan Akurasi Training NBC 2 Kelas

```
$Matrix_uji

predict        Lancar Tidak Lancar
  Lancar          131           62
  Tidak Lancar    357          538

$Akurasi
[1] 0.6148897

$Total_data
[1] 1088
```

f. Hasil Prediksi testing 25 teratas NBC 2 Kelas

```
 [1] Tidak Lancar Tidak Lancar Lancar       Tidak Lancar Tidak Lancar
 [6] Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar
[11] Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar
[16] Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar
[21] Tidak Lancar Lancar       Lancar       Tidak Lancar Lancar
Levels: Lancar Tidak Lancar
```

g. Hasil Prediksi dan Akurasi Testing NBC 2 Kelas

```
$Matrix_uji

predict        Lancar Tidak Lancar
  Lancar           25           13
  Tidak Lancar     96          136

$Akurasi
[1] 0.5962963

$Total_data
[1] 270
```

h. Data Training SVM 2 Kelas

	row.names	Jenjang	Fakultas	IPK	Jenis_Kelamin	Ketepatan_Studi	Organisasi	Y
1	1047	1	4	1	1	1	2	Lancar
2	245	0	0	2	1	1	1	Lancar
3	639	2	4	2	1	1	1	Lancar
4	1221	1	4	2	1	1	1	Lancar
5	217	1	6	2	0	1	0	Lancar
6	1245	1	6	1	0	1	2	Tidak Lancar
7	503	1	7	2	1	0	2	Tidak Lancar
8	1306	1	0	2	0	1	2	Lancar
9	261	1	6	1	0	0	4	Lancar
10	489	1	7	2	0	1	3	Lancar
11	187	1	1	0	0	1	3	Lancar
12	112	2	4	2	0	1	2	Lancar
13	659	1	3	2	1	1	2	Tidak Lancar
14	1150	2	4	2	1	1	1	Lancar
15	621	1	7	1	1	0	4	Tidak Lancar
16	785	1	5	1	1	0	1	Lancar
17	1206	1	4	2	0	0	4	Tidak Lancar
18	989	1	5	1	1	1	3	Lancar
19	310	2	4	2	1	1	0	Lancar
20	483	1	6	2	0	1	0	Lancar
21	1292	1	7	1	0	0	4	Lancar
22	921	1	1	1	0	1	3	Tidak Lancar
23	1252	1	6	1	1	1	3	Lancar
24	1181	1	5	1	1	1	1	Tidak Lancar
25	413	1	3	0	0	1	2	Tidak Lancar

i. Data Testing SVM 2 Kelas

	row.names	Jenjang	Fakultas	IPK	Jenis_Kelamin	Ketepatan_Studi	Organisasi	Y
1	8	1	7	1	1	1	4	Tidak Lancar
2	10	0	0	1	0	1	2	Tidak Lancar
3	11	2	4	2	1	0	3	Tidak Lancar
4	12	1	3	0	0	0	2	Tidak Lancar
5	18	1	6	0	0	0	2	Lancar
6	32	1	0	2	0	1	3	Lancar
7	37	1	7	1	0	0	3	Tidak Lancar
8	38	1	7	1	0	0	1	Tidak Lancar
9	53	1	7	0	0	1	2	Tidak Lancar
10	55	1	6	1	0	1	2	Tidak Lancar
11	65	1	4	2	1	1	1	Tidak Lancar
12	68	1	7	1	0	1	2	Lancar
13	69	1	0	1	1	1	0	Tidak Lancar
14	72	1	7	1	0	1	1	Tidak Lancar
15	76	1	0	2	1	1	3	Tidak Lancar
16	77	1	7	1	0	0	0	Tidak Lancar
17	78	1	4	1	1	1	1	Lancar
18	81	1	6	1	0	0	1	Tidak Lancar
19	83	1	7	2	1	1	2	Tidak Lancar
20	85	1	7	1	1	1	3	Lancar
21	94	1	2	2	1	1	2	Tidak Lancar
22	95	1	2	1	1	1	1	Tidak Lancar
23	98	1	5	1	1	1	1	Tidak Lancar
24	101	2	4	1	1	1	2	Tidak Lancar
25	109	2	4	2	1	1	2	Lancar

j. Hasil Model SVM 2 Kelas

```
Call:
svm(formula = Y ~ ., data = training, type = "C-classification", kernel = "polynomial", cost = 0.1,
    d = 1)


Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  polynomial
       cost:  0.1
     degree:  1
      gamma:  0.1666667
     coef.0:  0

Number of Support Vectors:  955

 ( 477 478 )


Number of Classes:  2

Levels:
 Lancar Tidak Lancar
```

k. Hasil Prediksi training 25 teratas SVM 2 Kelas

```
        1047          245          639         1221          217         1245
Tidak Lancar Tidak Lancar       Lancar Tidak Lancar Tidak Lancar Tidak Lancar
         503         1306          261          489          187          112
Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar       Lancar
         659         1150          621          785         1206          989
Tidak Lancar       Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar
         310          483         1292          921         1252         1181
      Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar
         413
Tidak Lancar
Levels: Lancar Tidak Lancar
```

l. Hasil Prediksi training dan Akurasi SVM 2 Kelas

```
$Matrix_Uji

pred2          Lancar Tidak Lancar
  Lancar           96           27
  Tidak Lancar    408          555

$Akurasi
[1] 0.5994475

$Total_data
[1] 1086
```

m. Hasil Tuning SVM 2 Kelas

```
Parameter tuning of 'svm':

- sampling method: 10-fold cross validation

- best parameters:
 cost d
    1 1

- best performance: 0.3130952
```

n. Hasil Model Testing SVM 2 Kelas

```
Call:
svm(formula = Y ~ ., data = testing, type = "C-classification", kernel = "polynomial", cost = 0.1,
    d = 1)


Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  polynomial
       cost:  0.1
     degree:  1
      gamma:  0.1666667
     coef.0:  0

Number of Support Vectors:  213
```

o. Hasil Prediksi testing 25 teratas SVM 2 Kelas

```
           8           10           11           12           18           32
Tidak Lancar Tidak Lancar       Lancar Tidak Lancar Tidak Lancar Tidak Lancar
          37           38           53           55           65           68
Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar
          69           72           76           77           78           81
Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar
          83           85           94           95           98          101
Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar Tidak Lancar       Lancar
         109
      Lancar
Levels: Lancar Tidak Lancar
```

p. Hasil Prediksi dan Akurasi SVM 2 Kelas

```
$Matrix_Uji

predsvmtest    Lancar Tidak Lancar
  Lancar           26            6
  Tidak Lancar     79          161

$Akurasi
[1] 0.6875

$Total_data
[1] 272
```

q. Nilai Kappa SVM 2 Kelas

```
$diag
[1] 0.6875

$kappa
[1] 0.2430592

$rand
[1] 0.5687269

$crand
[1] 0.1123664
```