

**PENERAPAN TEKNIK *DATA MINING* DALAM PREDIKSI MASA  
TUNGGU MAHASISWA MENGGUNAKAN METODE KLASIFIKASI**

**SKRIPSI**

Sebagai salah satu syarat untuk menyelesaikan studi  
di Program Studi Sistem Informasi S1



**Oleh:**

**Iqbal Malik Alfaruq**

**09031281823049**

**PROGRAM STUDI SISTEM INFORMASI**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS SRIWIJAYA**

**NOVEMBER 2022**

**HALAMAN PENGESAHAN**

**PENERAPAN TEKNIK *DATA MINING* DALAM PREDIKSI MASA  
TUNGGU MAHASISWA MENGGUNAKAN METODE KLASIFIKASI**

**SKRIPSI**

Sebagai salah satu syarat untuk menyelesaikan studi  
di Program Studi Sistem Informasi SI

Oleh:

**Iqbal Malik Alfaruq**

**09031281823049**

**Palembang, 15 November 2022**

**Mengetahui**

**Ketua Jurusan,**



**Endang Lestari Ruskan, M.T.**

**NIP 19781117 200604 2 001**

**Pembimbing**

A handwritten signature in black ink, consisting of several loops and a long horizontal stroke.

**Ken Ditha Tania, M.Kom.**

**NIP 19850718 201212 2 003**

## HALAMAN PERNYATAAN BEBAS PLAGIAT

Yang bertanda tangan di bawah ini:

Nama : Iqbal Malik Alfaruq

NIM : 09031281823049

Program Studi : Sistem Informasi

Judul Skripsi : Penerapan Teknik Data Mining Dalam Prediksi Masa Tunggu  
Mahasiswa Menggunakan Metode Klasifikasi

Hasil Pengecekan *Software iThenticate*/Turnitin: 18% (delapan belas persen)

Menyatakan bahwa laporan skripsi Saya merupakan hasil karya sendiri dan bukan penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan skripsi ini, maka Saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan berlaku.

Demikian pernyataan ini Saya buat dengan sebenarnya dan tidak ada unsur paksaan dari siapapun.

Palembang, 15 November 2022



Iqbal Malik Alfaruq

NIM 09031281823049

## HALAMAN PERSETUJUAN

Telah diuji dan lulus pada:

Hari : Jumat

Tanggal : 15 Agustus 2022

Tim Penguji :

1. Pembimbing : Ken Ditha Tania, M.Kom.
2. Ketua : Ari Wedhasmara, M.TI.
3. Penguji I : Dedy Kurniawan, M.Sc.
4. Penguji II : Ali Bardadi, M.Kom.



Mengetahui,

Ketua Jurusan Sistem Informasi



Endang Lestari Ruskan, S.Kom., M.T.

NIP 19781117 200604 2 001

## HALAMAN PERSEMBAHAN

**Motto:**

**“Hal yang salah jangan ditutupi, tapi harus diperbaiki”**

Iqbal Malik Alfaruq, 2022

Penulis mempersembahkan skripsi ini kepada:

1. Diri sendiri,
2. Orang tua, kakak laki-laki dan adik perempuan penulis,
3. Dosen pembimbing penulis,
4. Dosen penguji penulis,
5. Sahabat-sahabat penulis, serta
6. Universitas Sriwijaya.

# **PENERAPAN TEKNIK *DATA MINING* DALAM PREDIKSI MASA TUNGGU MAHASISWA MENGGUNAKAN METODE KLASIFIKASI**

Oleh :  
Iqbal Malik Alfaruq  
09031281823049

## **ABSTRAK**

Pengangguran adalah salah satu masalah krusial yang dihadapi oleh Indonesia. Terlebih pengangguran lulusan Sarjana dan Diploma yang tiap tahun memiliki angka yang besar. Berdasarkan data resmi yang diterbitkan oleh Badan Pusat Statistik (BPS) Indonesia tercatat pengangguran terbuka lulusan Sarjana dan Diploma sejak tahun 2019 sampai 2021 belum mencapai kurang dari 9 ratus ribu orang. Dari data tersebut digunakanlah dataset masa tunggu kerja setelah lulus dari tahun 2019 hingga 2021 yang telah didapatkan oleh *Career Development Center* Universitas Sriwijaya (CDC Unsri) dengan kegiatan *Tracer Study* yang melakukan kuesioner terhadap alumni Universitas Sriwijaya. Penelitian ini menggunakan metode *Cross Industry Standard Process for Data Mining* (CRISP-DM) karena memiliki tahap-tahap yang digunakan untuk data mining, masing-masing tahap mempunyai fungsinya tersendiri akan tetapi berhubungan dengan lainnya. Tahapannya berupa fase pemahaman bisnis, fase pemahaman data, fase persiapan data, fase pemodelan, fase evaluasi, dan fase penyebaran. Berdasarkan penelitian ini didapatkan variabel yang memiliki hubungan dengan masa tunggu adalah masa studi tahun, bahasa inggris, kemampuan analisis, kepemimpinan, dan menulis dokumen. Kemudian algoritma yang digunakan adalah *Random Forest* karena memiliki keunggulan dalam menghasilkan hasil klasifikasi yang bagus, dapat memproses data yang besar dan memberikan hasil yang kurang tersebar tetapi lebih inovatif. Dan menghasilkan data latih kombinasi tahun 2021, 2020 & 2019 memiliki akurasi yang sebesar 78.17% dan standar deviasi sebesar 1.85% yang tentu saja memenuhi kriteria tidak lebih dari 2%.

**Kata Kunci:** *Data Mining*, Klasifikasi, Masa Tunggu, CRISP-DM

# APPLICATION OF DATA MINING TECHNIQUES IN PREDICTING STUDENT WAITING PERIOD USING THE CLASSIFICATION METHOD

By:

**Iqbal Malik Alfaruq 0903128182.049**

Major of Information Systems, Faculty of Computer Science, Sriwijaya  
University

Email: iqbalma.00@gmail.com

## ABSTRACT

Unemployment is one of the crucial problems faced by Indonesia. Moreover, unemployment for Bachelor and Diploma graduates, which has a large number every year. Based on official data published by the Indonesian Central Bureau of Statistics (CBS), it is recorded that open unemployment for Bachelor and Diploma graduates from 2019 to 2021 has not reached less than 9 hundred thousand people. From this data, the waiting period dataset after graduating from 2019 to 2021 was used which was obtained by the Sriwijaya University Career Development Center (CDC Unsri) with a Tracer Study activity that conducted questionnaires on alumni of Sriwijaya University. This study uses the Cross Industry Standard Process for Data Mining (CRISP-DM) method because it has stages used for data mining, each stage has its own function but is related to the others. The stages are in the form of business understanding phase, data understanding phase, data preparation phase, modeling phase, evaluation phase, and deployment phase. Based on this study, it was found that the variables that had a relationship with the waiting period were years of study, English language skills, analytical skills, leadership, and writing documents. Then the algorithm used is Random Forest because it has advantages in producing good classification results, can process large data and provides results that are less scattered but more innovative. And producing combined training data for 2021, 2020 & 2019 has an accuracy of 78.17% and a standard deviation of 1.85% which of course meets the criteria of no more than 2%.

**Keywords:** *Data Mining*, Classification, Waiting Period, CRISP-DM

## KATA PENGANTAR



Puji dan syukur penulis panjatkan atas kehadiran Allah SWT karena berkat rahmat, karunia, serta hidayah-Nya, penulis dapat menyelesaikan skripsi dengan judul “Penerapan Teknik Data Mining Dalam Prediksi Masa Tunggu Mahasiswa Menggunakan Metode Klasifikasi” ini hingga selesai. Skripsi ini merupakan salah satu syarat untuk menyelesaikan studi penulis di program studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Sriwijaya.

Penulis menyadari bahwa skripsi ini tidak mungkin dapat diselesaikan tanpa adanya bantuan dari berbagai pihak. Maka dari itu, penulis ingin mengucapkan terima kasih kepada:

1. Allah SWT yang telah mencurahkan rahmat, karunia, serta hidayah-Nya sehingga dalam perjalanan prosesan sampai penulisan skripsi ini selalu diberikan kekuatan dan ketabahan.
2. Orang tua, kakak & adik penulis, Alm. Drs. Nasir Muchsin, Sudarti, S.Pd.SD., Ilham Ilahiya Ramadlan dan Izzah Syifaul Qulub yang selalu memberikan dorongan dan semangat kepada penulis untuk menyelesaikan studi Strata 1.
3. Bapak Jaidan Jauhari, S.Pd., M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Ibu Endang Lestari Ruskan, S.Kom., M.T., selaku Kepala Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.



5. Ibu Ken Ditha Tania, M.Kom., selaku Dosen Pembimbing Akademik, Dosen Pembimbing Kerja Praktik dan sebagai Dosen Pembimbing Skripsi yang selalu membimbing penulis mulai dari tahun pertama perkuliahan hingga sampai penulisan skripsi ini selesai.
6. Bapak Ari Wedhasmara, M.TI., selaku Dosen Ketua Penguji, Bapak Dedy Kurniawan, M.Sc., selaku Dosen Penguji I, dan Bapak Ali Bardadi, M.Kom., selaku Dosen Penguji II.
7. Mbak Rifka, selaku Admin Jurusan Sistem Informasi Bilingual yang membantu dan menginformasikan segala kebutuhan selama masa perkuliahan penulis.
8. Ibu Dwi Rosa Indah, S.T., M.T., selaku Kepala Laboratorium Struktur Data dan Sistem Informasi Akuntansi serta Royan Dwi Saputra, Rofiqul Rahman Ramadhan dan Dina Primalaya selaku Asisten Laboratorium Struktur Data dan Sistem Informasi Akuntansi kampus Bukit, yang menemani proses dan sampai penulisan skripsi penulis selesai.
9. Sahabat penulis, Biancha Anisyah Lestari, Rijal Abdul Hadi Harahap dan Muhammad Risky Saputra yang selalu menemani di setiap proses dan sampai penulisan skripsi penulis selesai.
10. Teman satu pembimbingan, Muhammad Frido yang selalu memberikan masukan dan saran selama proses dan sampai penulisan skripsi penulis selesai.
11. Teman-teman seperjuangan penulis, Sistem Informasi Bilingual A 2018 yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa skripsi ini masih belum sempurna. Maka dari itu, penulis dengan tangan terbuka menerima semua bentuk kritik dan saran yang membangun dari pembaca untuk kebaikan penulisan penulis di masa yang akan datang. Penulis juga berharap skripsi ini dapat memberikan manfaat kepada pembaca, baik dari sisi teoritis maupun praktis.

Palembang, 15 November 2022

Penulis



Iqbal Malik Alfaruq

NIM.09031281823049

## DAFTAR ISI

HALAMAN JUDUL .....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN BEBAS PLAGIAT .....	iii
HALAMAN PERSETUJUAN .....	iv
HALAMAN PERSEMBAHAN .....	v
ABSTRAK .....	vi
KATA PENGANTAR.....	viii
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL .....	xv
DAFTAR LAMPIRAN .....	xvi
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian .....	4
1.4 Manfaat Penelitian .....	5
1.5 Batasan Masalah.....	5
<b>BAB II TINJAUAN PUSATAKA.....</b>	<b>7</b>
2.1 Penelitian Sebelumnya .....	7
2.2 <i>Tracer Study</i> .....	8
2.3 <i>Data Mining</i> .....	9
2.3.1 Definisi <i>Data Mining</i> .....	10
2.3.2 Proses <i>Data Mining</i> .....	11
2.3.3 Arsitektur Sistem <i>Data Mining</i> .....	13
2.4 Model <i>Waterfall</i> .....	14
2.5 Cross Industry Standard Process for Data Mining (CRISP-DM) ....	16
2.5 Metode Klasifikasi.....	17
2.6 <i>Random Forest</i> .....	18
2.7 <i>10 Fold-Cross Validation</i> .....	19
a. <i>Confusion Matrix</i> .....	20
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>22</b>
3.1 Objek Penelitian .....	22
3.2 Pengumpulan Data .....	22
3.2.1 Metode Pengumpulan Data.....	22
3.2.2 Jenis Data dan Sumber Data.....	23
3.3 Kerangka Penelitian .....	24
3.4 Fase Analisis Kebutuhan .....	25
3.4.1 Kebutuhan Sistem .....	25
3.4.2 Perancangan Interaksi Manusia Komputer .....	26

3.5	Fase Perancangan .....	38
3.5.1	Metode CRISP-DM .....	38
3.5.2	Rancangan <i>Interface</i> .....	57
3.6	Fase Implementasi .....	66
3.7	Fase Pengujian .....	67
3.8	Fase Pemeliharaan .....	67
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>		<b>68</b>
4.1	Hasil .....	68
4.2	Pembahasan .....	69
4.2.1	Halaman <i>Login</i> .....	69
4.2.2	Halaman Registrasi Akun .....	70
4.2.3	Halaman Formulir Prediksi Mahasiswa .....	71
4.2.4	Halaman <i>Data Training</i> .....	73
4.2.5	Halaman Akurasi Data .....	74
4.2.6	Halaman Data Mahasiswa .....	75
4.2.7	Halaman Akun Mahasiswa .....	76
4.2.8	Halaman Verifikasi Akun Mahasiswa .....	78
4.3	Pengujian Sistem Aplikasi Prediksi Masa Tunggu Mahasiswa .....	79
<b>BAB V KESIMPULAN DAN SARAN .....</b>		<b>85</b>
5.1	Kesimpulan .....	85
5.2	Saran .....	86
<b>DAFTAR PUSTAKA .....</b>		<b>87</b>
<b>LAMPIRAN .....</b>		<b>A-1</b>

## DAFTAR GAMBAR

Gambar 1.1 Jumlab Pengangguran Terbuka di Indonesia Periode 2019-2021 Bulan Februari dan Agustus (dalam ratusan ribu jiwa) (Sumber : Badan Pusat Statistik, n.d.-a).....	1
Gambar 1.2 Proses Model <i>Waterfall</i> (Sumber : Heriyanti & Ishak, 2020).....	14
Gambar 2.1 Model CRISP-DM (Sumber : Mulaab, 2021) .....	16
Gambar 2.2 Ilustrasi Random Forest .....	18
Gambar 2.3 Ilustrasi Confusion Matrix .....	21
Gambar 3.1 Diagram Kerangka Penelitian .....	24
Gambar 3.2 <i>Use Case Diagram</i> Sistem Prediksi .....	28
Gambar 3.3 <i>Raw Data Tracer Study</i> .....	40
Gambar 3.4 Menu Korelasi.....	43
Gambar 3.5 Mengatur Setelan Korelasi .....	43
Gambar 3.6 Hasil Analisis Korelasi .....	46
Gambar 3.7 Jumlah Data Latih Berdasarkan Masa Tunggu .....	49
Gambar 3.8 Tabel Penyesuaian Sampel .....	50
Gambar 3.9 Total Data Latih Sampel .....	51
Gambar 3.10 Pemodelan Menggunakan Cross Validation .....	52
Gambar 3.11 Pemodelan didalam CV Menggunakan Random Forest .....	53
Gambar 3.12 Pemodelan Prediksi Masa Tunggu Menggunakan <i>Random Forest</i> Pada Sistem .....	54
Gambar 3.13 Pemodelan Cek Akurasi <i>Data Training</i> Pada Sistem .....	55
Gambar 3.14 Evaluasi Akurasi Data.....	56
Gambar 3.15 Evaluasi Standar Deviasi .....	56
Gambar 3.16 Rancangan Halaman <i>Login</i> .....	58
Gambar 3.17 Rancangan Halaman Registrasi .....	59
Gambar 3.18 Rancangan Halaman Formulir Prediksi Mahasiswa .....	59
Gambar 3.19 Rancangan Halaman Prediksi Mahasiswa.....	60
Gambar 3.20 Rancangan Halaman Formulir Sunting Prediksi Mahasiswa..	60
Gambar 3.21 Rancangan Halaman <i>Data Training</i> .....	61
Gambar 3.22 Rancangan Halaman Perbandingan <i>Data Training</i> .....	62
Gambar 3.23 Rancangan Halaman Data Mahasiswa.....	63
Gambar 3.24 Rancangan Halaman Detail Data Mahasiswa.....	64
Gambar 3.25 Rancangan Halaman Akun Mahasiswa .....	64
Gambar 3.26 Rancangan Halaman Verifikasi Akun Mahasiswa.....	65
Gambar 3.27 Rancangan Halaman Detail Verifikasi Mahasiswa. ....	66
Gambar 4.1 Halaman <i>Login</i> .....	69
Gambar 4.2 Pesan Gagal <i>Login</i> .....	70
Gambar 4.3 Halaman Registrasi Akun.....	70
Gambar 4.4 Pesan Gagal Registrasi.....	71
Gambar 4.5 Halaman Formulir Prediksi Mahasiswa .....	71
Gambar 4.6 Pesan Galat Formulir Prediksi Kosong.....	72
Gambar 4.7 Halaman Prediksi Mahasiswa .....	72
Gambar 4.8 Halaman Formulir Sunting Prediksi Mahasiswa .....	73
Gambar 4.9 Halaman <i>Data Training</i> .....	73

<b>Gambar 4.10</b>	<b>Formulir Unggah <i>Data Training</i></b> .....	<b>74</b>
<b>Gambar 4.11</b>	<b>Halaman Akurasi Data</b> .....	<b>74</b>
<b>Gambar 4.12</b>	<b>Halaman Data Mahasiswa</b> .....	<b>75</b>
<b>Gambar 4.13</b>	<b>Halaman Detail Mahasiswa</b> .....	<b>76</b>
<b>Gambar 4.14</b>	<b>Halaman Akun Mahasiswa</b> .....	<b>76</b>
<b>Gambar 4.15</b>	<b>Halaman Detail Akun Mahasiswa</b> .....	<b>77</b>
<b>Gambar 4.16</b>	<b>Pesan Ketika Hapus Akun</b> .....	<b>77</b>
<b>Gambar 4.17</b>	<b>Halaman Verifikasi Akun Mahasiswa</b> .....	<b>78</b>
<b>Gambar 4.18</b>	<b>Halaman Detail Verifikasi Akun Mahasiswa</b> .....	<b>78</b>
<b>Gambar 4.19</b>	<b>Pesan Ketika Hapus Akun yang Diverifikasi</b> .....	<b>79</b>

## DAFTAR TABEL

Tabel 2.1 Ilustrasi 10 K Fold Cross V .....	20
Tabel 3.1 Analisis Kebutuhan Fungsional .....	26
Tabel 3.2 Analisis Kebutuhan Non Fungsional .....	26
Tabel 3.3 Definisi <i>Use Case</i> .....	28
Tabel 3.4 Skenario Melihat Halaman Beranda .....	30
Tabel 3.5 Skenario <i>Login</i> .....	30
Tabel 3.6 Skenario Melakukan Registari Akun Mahasiswa.....	30
Tabel 3.7 Skenario Mengisi Form Prediksi .....	31
Tabel 3.8 Skenario Melihat Hasil Prediksi .....	32
Tabel 3.9 Skenario Menyunting Prediksi .....	32
Tabel 3.10 Skenario Memvalidasi Akun Mahasiswa.....	33
Tabel 3.11 Skenario Menambahkan <i>Data Training</i> .....	33
Tabel 3.12 Skenario Membandingkan <i>Data Training</i> .....	34
Tabel 3.13 Skenario Melihat Pedoman Memperbarui <i>Data Training</i> .....	34
Tabel 3.14 Skenario Melihat Daftar Prediksi Mahasiswa.....	35
Tabel 3.15 Skenario Melihat Detail Prediksi Mahasiswa.....	35
Tabel 3.16 Skenario Melihat Daftar Akun Mahasiswa .....	36
Tabel 3.17 Skenario Menghapus Akun Mahasiswa.....	36
Tabel 3.18 Dampak Tracer Study.....	39
Tabel 3.19 Partisi Nilai Variabel Prediksi .....	47
Tabel 4.1 Pengujian Kebutuhan Sistem.....	79
Tabel 4.2 Pengujian Operasional Sistem .....	80

## DAFTAR LAMPIRAN

<b>Lampiran 1. Surat Kesediaan Membimbing .....</b>	<b>A-1</b>
<b>Lampiran 2. Kartu Konsultasi.....</b>	<b>B-1</b>
<b>Lampiran 3. Lembar Rekomendasi Ujian Akhir.....</b>	<b>C-1</b>
<b>Lampiran 4. Form Perbaikan Ujian Komprehensif .....</b>	<b>D-1</b>
<b>Lampiran 5. Hasil Pengecekan Turnitin.....</b>	<b>E-1</b>

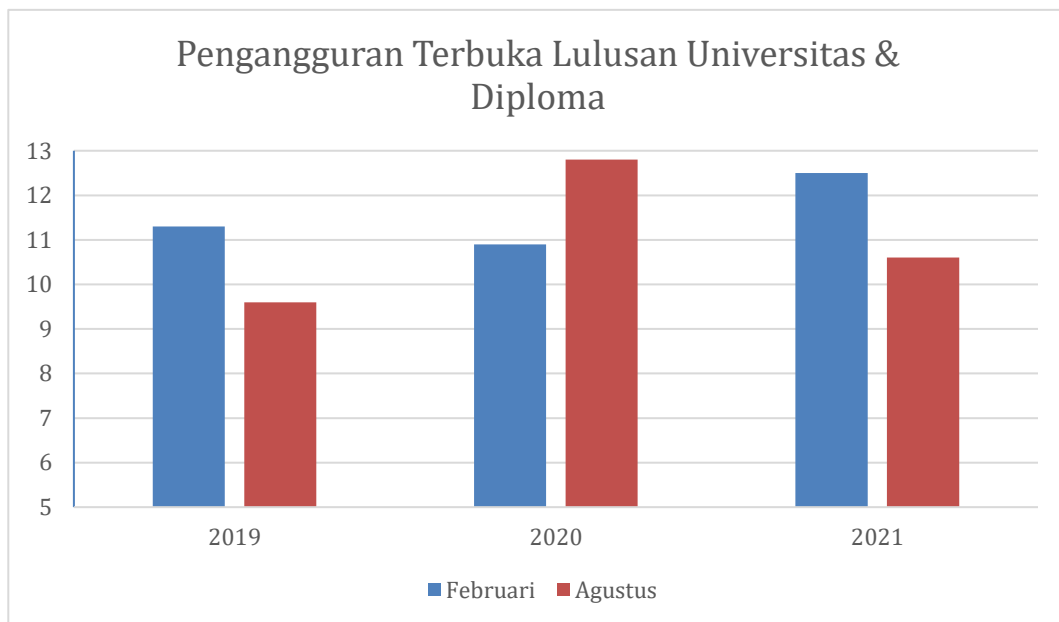


# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pengangguran adalah masalah yang sulit untuk ditangani yang pasti ada di setiap negara berkembang karena hal ini bisa mengakibatkan ekonomi yang kurang sehat dan intensitas kriminal menjadi naik (Astriani & Nooraeni, 2020). Penganggur terbuka, terdiri dari: Mereka yang tak punya pekerjaan dan mencari pekerjaan; mereka yang tak punya pekerjaan dan mempersiapkan usaha; Mereka yang tak punya pekerjaan dan tidak mencari pekerjaan, karena merasa tidak mungkin mendapatkan pekerjaan (Badan Pusat Statistik, n.d.-b).



**Gambar 1.1** Jumlah Pengangguran Terbuka di Indonesia Periode 2019-2021  
Bulan Februari dan Agustus (dalam ratusan ribu jiwa)

(Sumber : Badan Pusat Statistik, n.d.-a)

Berdasarkan data resmi yang diterbiskan oleh Badan Pusat Statistik, jumlah pengangguran terbuka lulusan universitas dan diploma sejak tahun 2019 sampai dengan tahun 2021 belum pernah mencapai kurang dari 9 ratus ribu orang. Hal tersebut bukan tanpa sebab, melainkan ada beberapa faktor yang perlu diamati.

Adapun faktor yang mempengaruhi masa tunggu kerja menurut Asroni et al. (2018) “atribut yang mempengaruhi masa tunggu yaitu tahun lulus, ipk, jenis pekerjaan, fakultas dan angkatan.” Serta menurut Khoirudin et al.(2020) “variabel yang digunakan untuk prediksi adalah IPK, jumlah alumni, lama waktu study, kelamin, kemampuan bahasa inggris dan kemampuan *softskill*”. Dan penelitian yang dilakukan oleh Rezkika et al.(2021) menyatakan bahwa “atribut yang dipilih adalah jenis kelamin, angkatan, tahun lulus, IPK, lama studi, riwayat organisasi/kepanitiaan, pengalaman magang/kerja, memiliki sertifikasi keahlian, lama waktu mendapatkan pekerjaan, bidang pekerjaan, jenis perusahaan, kota tempat bekerja, dan kesesuaian bidang pekerjaan dengan TI” dan juga menurut Palupi et al. (2018) “kriteria kualifikasi calon karyawan berupa pendidikan terakhir, pengalaman kerja, umur, status, kemampuan komputer, dan fasih bahasa Inggris”.

Kemampuan diri adalah hal yang sangat positif untuk dipunyai dan berguna agar tetap fokus, terhubung serta mandiri. Kemampuan diri ini sama-sama bermanfaat baik untuk kehidupan pribadi maupun dunia kerja. Itulah kenapa hal ini perlu dikembangkan karena pengusaha tentu mencari orang yang memiliki keterampilan saat merekrut karyawan baru (Atiqur et al., 2021).

Masa tunggu kerja adalah kondisi dimana berapa lama waktu yang diperlukan untuk mendapatkan pekerjaan pertama yang dihitung setelah lulus menempuh pendidikan. Dan persentasi alumni yang langsung bekerja menjadi salah

satu variabel penilaian akreditasi untuk perguruan tinggi serta program studi (Yulmiati et al., 2021).

Universitas Sriwijaya adalah salah satu perguruan tinggi negeri yang ada di Sumatera Selatan. Di Universitas Sriwijaya Mengenai banyak.Mahasiswa aktif serta alumni. Oleh karena itu, pihak Universitas Sriwijaya mempunyai data alumni mahasiswa yang bisa dimanfaatkan dan diolah untuk mendapatkan pengetahuan baru dan juga kaitan tertentu yang menciptakan informasi yang sangat berguna.

Pengaplikasian datang mining dengan menggunakan data alumni yang bisa membantu universitas Sriwijaya untuk mencari hubungan tingkat prestasi akademik terhadap masa tunggu kerja serta mendapatkan kaitan data yang terkandung di dalam setiap basis data. Menurut Nastuti (2019) “*Data mining* ialah suatu cara untuk menemukan pola dan hubungan yang berarti dengan memeriksa dalam sekumpulan banyak data yang tersimpan dalam penyimpanan data dengan menggunakan teknik pengenalan pola”

Data alumni yang banyak ini tentu dapat diolah untuk mendapatkan pengetahuan baru yang bermanfaat untuk pihak Universitas Sriwijaya. Hal ini sejalan dengan pengertian untuk menemukan pola dan hubungan yang berarti didalam penyimpanan data.

Pada penelitian kali ini bertujuan untuk merancang sistem berbasis web yang dapat memprediksi masa tunggu berdasarkan prestasi akademik dan keterampilan diri dengan metode klasifikasi. Data yang digunakan merupakan data alumni yang telah dikumpulkan UPT Pusat Pengembangan Karakter dan Pengelolaan Alumni Universitas Sriwijaya melalui pengisian *tracer study* Universitas Sriwijaya tahun 2021 oleh alumni yang lulus angkatan 2019. Kuesioner

ini memiliki banyak pertanyaan, dari begitu banyak pertanyaan, peneliti hanya berfokus pada prestasi akademik, potensi diri dan lama masa tunggu.

Berdasarkan uraian diatas penulis akan melakukan penelitian yang berjudul

**“Penerapan Teknik *Data Mining* Dalam Prediksi Masa Tunggu Mahasiswa  
Menggunakan Metode Klasifikasi”**

## **1.2 Rumusan Masalah**

Menurut latar belakang tersebut, maka penulis merumuskan masalah dari penelitian ini sebagai berikut :

1. Bagaimana prestasi akademik dan keterampilan diri berpengaruh terhadap masa lama tunggu.
2. Bagaimana hasil prediksi masa tunggu menggunakan metode klasifikasi.
3. Bagaimana cara menerapkan metode klasifikasi untuk memprediksi lama masa tunggu mahasiswa.

## **1.3 Tujuan Penelitian**

Tujuan yang ingin penulis hasilkan dari penelitian ini adalah :

1. Mengetahui bagaimana prestasi akademik dan keterampilan diri bisa berpengaruh terhadap masa tunggu.
2. Mengetahui hasil yang didapatkan dari penerapan metode klasifikasi untuk memprediksi lama masa tunggu.
3. Menghasilkan sistem aplikasi berbasis web yang bisa memprediksi lama masa tunggu berdasarkan prestasi akademik dan keterampilan diri.

#### **1.4 Manfaat Penelitian**

Hasil dari penelitian ini memberikan manfaat sebagai berikut :

1. Pihak Universitas Sriwijaya dapat melihat bagaimana prestasi akademik berpengaruh terhadap lama masa tunggu menggunakan sistem aplikasi berbasis web yang telah dibuat.
2. Keluaran yang dihasilkan dari penelitian ini diharapkan bisa menjadi bahan acuan bagi pemegang keputusan untuk dapat memperbaiki proses belajar dan kurikulum Universitas Sriwijaya agar dapat mempersingkat masa tunggu alumni.
3. Diharapkan penelitian ini bisa menjadi referensi bagi penelitian serupa di masa yang akan datang nanti.

#### **1.5 Batasan Masalah**

Batasan masalah dari penelitian ini sebagai berikut :

1. Studi kasus yang dilakukan pada penelitian ini yaitu Universitas Sriwijaya.
2. Data yang dipakai ialah data alumni Universitas Sriwijaya yang berhasil dikumpulkan UPT Pusat Pengembangan Karir dan Pengelolaan Alumni Universitas Sriwijaya atau bisa disebut sebagai Career Development Center Universitas Sriwijaya (CDC Unsri).
3. Data yang digunakan adalah alumni yang mendapatkan pekerjaan setelah lulus.
4. Penelitian ini memakai metode klasifikasi.
5. Tahapan dalam penelitian ini memakai tahapan pada metode *Cross Industry Process for Data Mining (CRISP-DM)*.

## DAFTAR PUSTAKA

- Acceptable Standard Deviation (SD)* - LabCE.com, *Laboratory Continuing Education*. (n.d.). Retrieved September 26, 2022, from [https://www.labce.com/spg49741\\_acceptable\\_standard\\_deviation\\_sd.aspx](https://www.labce.com/spg49741_acceptable_standard_deviation_sd.aspx)
- Aldo, D., Syawitri, A., Alwendi, Darmansah, & Samosir, K. (2021). *Data Mining* (A. Primayeszky (Ed.); 1st ed.). ICM Publisher.
- Anjani, N. S. (2018). Pengaruh Prestasi Belajar, Masa Studi dan Keaktifan Berorganisasi Terhadap Masa Tunggu dan Relevansi Pekerjaan Lulusan Prodi Pendidikan Ekonomi Fakultas Ekonomi Universitas Negeri Yogyakarta. *Skripsi Fakultas Ekonomi, Universitas Negeri Yogyakarta*, 131–132.
- Asroni, A., Maharty Ali, N., & Riyadi, S. (2018). Perkiraan Masa Tunggu Alumni Mendapatkan Pekerjaan Menggunakan Metode Prediksi Data Mining Dengan Algoritma Naive Bayes Classifier. *Semesta Teknika*, 21(2), 189–197. <https://doi.org/10.18196/st.212225>
- Astriani, V., & Nooraeni, R. (2020). Determinan Pengangguran Lulusan Perguruan Tinggi Di Indonesia Tahun 2018. *Jurnal Pendidikan Ekonomi (JUPE)*, 8(1), 31–37. <https://doi.org/10.26740/jupe.v8n1.p31-37>
- Atiqur, M., Sarker, R., Hashim, J. B., Haque, A., & Juhdi, N. B. (2021). Nguồn về Soft Skills, đoạn research model. *JIBM) Journal of International Business and Management*, 4(4), 1–14. <https://rpajournals.com/jibmJournalHomepage:https://rpajournals.com/jibm>
- Badan Pusat Statistik. (n.d.-a). *Pengangguran Terbuka Menurut Pendidikan Tertinggi yang Ditamatkan*. Retrieved November 21, 2022, from <https://www.bps.go.id/indicator/6/674/1/pengangguran-terbuka-menurut-pendidikan-tertinggi-yang-ditamatkan.html>
- Badan Pusat Statistik. (n.d.-b). *Tenaga Kerja*. Retrieved November 21, 2022, from <https://www.bps.go.id/subject/6/tenaga-kerja.html>
- Barros, T. M., Neto, P. A. S., Silva, I., & Guedes, L. A. (2019). Predictive models for imbalanced data: A school dropout perspective. *Education Sciences*, 9(4). <https://doi.org/10.3390/educsci9040275>
- Chakure, A. (2019). *Random Forest Regression Along with its implementation in Python*. Medium. <https://medium.com/swlh/random-forest-and-its-implementation-71824ced454f>
- Corporate Finance Insitute. (2019). *Random Forest - Overview, Modeling Predictions, Advantages*. Corporate Finance Insitute. <https://corporatefinanceinstitute.com/resources/knowledge/other/random-forest/>
- Devella, S., Yohannes, Y., & Rahmawati, F. N. (2020). Implementasi Random Forest Untuk Klasifikasi Motif Songket Palembang Berdasarkan SIFT. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 7(2), 310–320. <https://doi.org/10.35957/jatisi.v7i2.289>
- Fahriya, K., & Nurhidayat, A. I. (2018). RANCANG BANGUN SIMAWA (SISTEM INFORMASI RUSUNAWA) BERBASIS WEBAPPLICATION MENGGUNAKAN FRAMEWORK LARAVEL. *Jurnal Manajemen Informatika*, 8.

- Ginantra, N. L. W. S. R., Arifah, F. N., Wijaya, A. H., Septarini, R. S., Ahmad, N., Ardiana, D. P. Y., Effendy, F., Iskandar, A., Hazriani, H., Sari, I. Y., Gustiana, Z., Prianto, C., Gustian, D., & Negara, E. S. (2021). *Data Mining dan Penerapan Algoritma*.
- Haristu, R. A., & Rosa, P. H. P. (2019). Penerapan Metode Random Forest untuk Prediksi Win Ratio Pemain Player Unknown Battleground. *MEANS (Media Informasi Analisa Dan Sistem)*, 4(2), 120–128. <https://doi.org/10.54367/means.v4i2.545>
- Hasanah, M. A., Soim, S., & Handayani, A. S. (2021). Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir. *Journal of Applied Informatics and Computing (JAIC)*, 5(2), 103–108. <http://jurnal.polibatam.ac.id/index.php/JAIC>
- Hasnain, M., Pasha, M. F., Ghani, I., Imran, M., Alzahrani, M. Y., & Budiarto, R. (2020). Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking. *IEEE Access*, 8, 90847–90861. <https://doi.org/10.1109/ACCESS.2020.2994222>
- He, L., Levine, R. A., Fan, J., Beemer, J., & Stronach, J. (2018). Random forest as a predictive analytics alternative to regression in institutional research. *Practical Assessment, Research and Evaluation*, 23(1), 1–16.
- Heriyanti, F., & Ishak, A. (2020). Design of logistics information system in the finished product warehouse with the waterfall method: Review literature. *IOP Conference Series: Materials Science and Engineering*, 801(1). <https://doi.org/10.1088/1757-899X/801/1/012100>
- Ismanto, E., & Novalia, M. (2021). Komparasi Kinerja Algoritma C4.5, Random Forest, dan Gradient Boosting untuk Klasifikasi Komoditas. *Techno.Com*, 20(3), 400–410. <https://doi.org/10.33633/tc.v20i3.4576>
- Jarmulska, B. (2022). Random forest versus logit models: Which offers better early warning of fiscal stress? *Journal of Forecasting*, 41(3), 455–490. <https://doi.org/10.1002/for.2806>
- Khoirudin, K., Hadi, S., & Nugroho, A. (2020). Analisa dan Penerapan Metode Neural Networks Dalam Mengidentifikasi Faktor-Faktor Masa Tunggu Kerja Lulusan. *Jurnal Pengembangan Rekayasa Dan Teknologi*, 16(1), 17. <https://doi.org/10.26623/jprt.v16i1.2399>
- Leonardo, R., Pratama, J., & Chrisnatalis, C. (2020). Perbandingan Metode Random Forest Dan Naïve Bayes Dalam Prediksi Keberhasilan Klien Telemarketing. *Jurnal Teknologi Dan Ilmu Komputer Prima (Jutikomp)*, 3(2), 1–5.
- Madison Schott. (2019, April 25). *Random Forest Algorithm for Machine Learning*. Medium. <https://medium.com/capital-one-tech/random-forest-algorithm-for-machine-learning-c4b2c8cc9feb>
- Mulaab. (2021). *Data Mining : Konsep dan Aplikasi* (Tim MNC Publishing (Ed.); 1st ed.). Media Nusa Creative.
- Nastuti, A. (2019). Amelia Nastuti 1 ) , Syaiful Zuhri Harahap 2 ). *Teknik Data Mining Untuk Penentuan Paket Hemat Sembako Dan Kebutuhan Harian Dengan Menggunakan Algoritma Fp-Growth*, 7(3), 111–119.
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal*

- Sains Komputer & Informatika (J-Sakti)*, 5(2), 697–711.
- Palupi, S., Lailiyah, S., & Sihotang, V. (2018). Sistem Pendukung Keputusan Penerimaan Karyawan Pada Pt. Suryaintan Tri Lestari Dengan Metode Ahp Berbasis Web. *Sebatik*, 11(1), 25–31. <https://doi.org/10.46984/sebatik.v11i1.97>
- Pangestu, S. Y., Astuti, Y., & Farida, L. D. (2019). Algoritma Support Vector Machine untuk Klasifikasi Sikap Politik terhadap Partai Politik Indonesia. *Jurnal Mantik Penusa*, 3(1), 236–241.
- Patel, H., Singh Rajput, D., Thippa Reddy, G., Iwendi, C., Kashif Bashir, A., & Jo, O. (2020). A review on classification of imbalanced data for wireless sensor networks. *International Journal of Distributed Sensor Networks*, 16(4). <https://doi.org/10.1177/1550147720916404>
- Pradana, A. W., & Hayaty, M. (2019). The Effect of Stemming and Removal of Stopwords on the Accuracy of Sentiment Analysis on Indonesian-language Texts. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 4(3), 375–380. <https://doi.org/10.22219/kinetik.v4i4.912>
- Prasetya, A. F., Sintia, S., & Putri, U. L. D. (2022). Perancangan Aplikasi Rental Mobil Menggunakan Diagram UML (Unified Modelling Language). *Jurnal Ilmiah Komputer ...*, 1(1), 14–18.
- Rezkika, F., Nurina Sari, B., Susilo Yuda Irawan, A., Studi Teknik Informatika, P., Singaperbangsa Karawang, U., Ronggo Waluyo, J. H., Jambe Timur, T., & Telp, K. (2021). Klasifikasi Masa Tunggu Alumni Untuk Mendapatkan Pekerjaan Berdasarkan Kompetensi Menggunakan Algoritma C4.5 (Studi Kasus : Fasilkom Unsika). *Progresif: Jurnal Ilmiah Komputer*, 17(2), 95–106. <http://ojs.stmik-banjarbaru.ac.id/index.php/progresif/article/view/652>
- Siburian, V. W., & Mulyana, I. E. (2018). Prediksi Harga Ponsel Menggunakan Metode Random Forest. *Prosiding Annual Research Seminar 2018*, 144–147.
- Yulmiati, Y., Rahmi H. G., I., & Yanuar, F. (2021). Klasifikasi Masa Tunggu Kerja Lulusan S1 Matematika Unand Menggunakan Naive Bayes Classifier. *Jurnal Riset Dan Aplikasi Matematika (JRAM)*, 5(1), 29. <https://doi.org/10.26740/jram.v5n1.p29-37>
- Zuffo, L. T., DeLima, R. O., & Lübberstedt, T. (2022). Combining datasets for maize root seedling traits increases the power of GWAS and genomic prediction accuracies. *Journal of Experimental Botany*. <https://doi.org/10.1093/jxb/erac236>