

**FEATURE SELECTION PADA ALGORITMA SUPPORT  
VECTOR MACHINE DAN RANDOM FOREST UNTUK  
KLASIFIKASI PENERIMA PROGRAM INDONESIA PINTAR**

**TUGAS AKHIR**

Diajukan Sebagai Salah Satu Syarat  
untuk Memperoleh Gelar Sarjana Komputer pada  
Program Studi Sistem Informasi

Oleh:



**NANDA TRY LUCHIA**

**12050320445**



**UIN SUSKA RIAU**  
**FAKULTAS SAINS DAN TEKNOLOGI**  
**UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU**  
**PEKANBARU**  
**2024**

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



**LEMBAR PERSETUJUAN**

***FEATURE SELECTION PADA ALGORITMA SUPPORT  
VECTOR MACHINE DAN RANDOM FOREST UNTUK  
KLASIFIKASI PENERIMA PROGRAM INDONESIA PINTAR***

**TUGAS AKHIR**

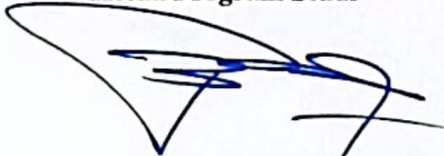
Oleh:

**NANDA TRY LUCHIA**

**12050320445**

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir  
di Pekanbaru, pada tanggal 20 Mei 2024

**Ketua Program Studi**



**Eki Saputra, S.Kom., M.Kom.**  
**NIP. 198307162011011008**

**Pembimbing**



**Mustakim, ST., M.Kom.**  
**NIK. 130511023**

**LEMBAR PENGESAHAN**

**FEATURE SELECTION PADA ALGORITMA SUPPORT  
VECTOR MACHINE DAN RANDOM FOREST UNTUK  
KLASIFIKASI PENERIMA PROGRAM INDONESIA PINTAR**

**TUGAS AKHIR**

Oleh:

**NANDA TRY LUCHIA**

**12050320445**

Telah dipertahankan di depan sidang dewan penguji  
sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer  
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau  
di Pekanbaru, pada tanggal 30 April 2024

Pekanbaru, 30 April 2024

Mengesahkan,

**Ketua Program Studi**



**Eki Saputra, S.Kom., M.Kom.**

**NIP. 198307162011011008**



**Dekan**  
**Dr. Harjono, M.Pd.**  
**NIP. 196403011992031003**

**DEWAN PENGUJI:**

**Ketua : Arif Marsal, Lc., MA.**

**Sekretaris : Mustakim, ST., M.Kom.**

**Anggota 1 : Fitriani Muttakin, S.Kom., M.Cs.**

**Anggota 2 : Idria Maita, S.Kom., M.Sc.**



## LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum, dengan ketentuan bahwa hak cipta ada pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan atas izin penulis dan harus dilakukan mengikuti kaedah dan kebiasaan ilmiah serta menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin tertulis dari Dekan Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan dapat meminjamkan Tugas Akhir ini untuk anggotanya dengan mengisi nama, tanda peminjaman dan tanggal pinjam pada *form* peminjaman.

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diuraikan dalam naskah ini dan disebutkan di dalam daftar pustaka.

Pekanbaru, 30 April 2024

Yang membuat pernyataan,

**NANDA TRY LUCHIA**

**NIM. 12050320445**

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LEMBAR PERSEMBAHAN

*Assalamu 'alaikum Warahmatullahi Wabarakatuh*

*Alhamdulillah Rabbil 'Alamin*, segala puji bagi Allah *Subhanahu Wa Ta'ala*

sebagai bentuk rasa syukur atas segala nikmat yang telah diberikan tanpa ada kekurangan sedikitpun. Shalawat beserta salam tak lupa pula kita ucapkan kepada Nabi Muhammad *Shallallahu 'Alaihi Wa Sallam* dengan mengucapkan "*Allahumma Sholli 'ala Sayyidina Muhammad Wa 'ala Ali Sayyidina Muhammad*". Semoga kita semua selalu senantiasa mendapat syafaat-Nya di dunia maupun di akhirat, *aamiin ya rabbal'aalamiin*.

Saya persembahkan hadiah istimewa karya kecil ini sebagai salah satu bentuk bakti, rasa terima kasih, dan hormat kepada Ayah dan Ibu tercinta. Terima kasih yang tak terhingga karena telah merawat dan membesarkan saya dengan setulus hati dan penuh perjuangan hingga saya bisa sampai tahap saat ini. Berkat doa dan kasih sayangmu, anakmu telah berhasil memperoleh gelar sarjana seperti yang engkau harapkan. Tiada apapun di dunia ini yang dapat membalas semua jasa-jasa dan pengorbananmu. Untuk itu saya anakmu ini selalu mendoakan yang terbaik untuk Ayah dan Ibu agar bahagia dunia dan akhirat serta diberikan tempat istimewa di sisi-Nya kelak. Dan pastinya saya juga berterima kasih yang tak terhingga kepada saudara kandung tercinta saya yaitu Abang dan Kakak yang telah memberikan saya pelajaran dan pemahaman mengenai indahnya kehidupan yang damai sebagai saudara.

Kepada Bapak dan Ibu Dosen Program Studi Sistem Informasi Universitas Islam Negeri Sultan Syarif Kasim Riau yang telah memberikan ilmu pengetahuan, pengalaman, dan kebaikan selama perkuliahan, saya ucapkan terima kasih banyak dan semoga menjadi amal jariyah. Aamiin.

Untuk sahabat terdekat yang tidak bisa saya sebutkan satu persatu dan pastinya juga teman-teman seperjuangan, terima kasih berkat kalian masa perkuliahan menjadi lebih bermakna semoga dimasa mendatang kita bisa bertemu lagi dalam keadaan yang lebih baik.

*Wassalamu 'alaikum Warahmatullahi Wabarakatuh*

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## KATA PENGANTAR

*Assalamu 'alaikum Warahmatullahi Wabarakatuh*

*Alhamdulillah Rabbil 'Alamin*, bersyukur kehadiran Allah *Subhanahu Wa Ta'ala* atas segala rahmat dan karunia-Nya sehingga peneliti dapat menyelesaikan Tugas Akhir ini. Tidak lupa sholawat beriringan salam selalu turunkan untuk Nabi Muhammad *Shallallahu 'Alaihi Wa Sallam* dengan melantunkan *Allahumma Sholli 'ala Sayyidina Muhammad Wa 'ala Alihi Muhammad*. Tugas Akhir ini dibuat sebagai salah satu syarat untuk mendapatkan gelar Sarjana Komputer di Program Studi Sistem Informasi Universitas Islam Negeri Sultan Syarif Kasim Riau.

Banyak orang yang telah berperan dalam mendukung dan membimbing peneliti pada proses penelitian dan penulisan Tugas Akhir ini. Terkhusus orang tua yang tidak pernah berhenti mendukung dan berdoa demi kesuksesan peneliti. Maka dari itu, ungkapan terima kasih juga peneliti ucapkan kepada:

1. Bapak Prof. Dr. Hairunas, M.Ag sebagai Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd sebagai Dekan Fakultas Sains dan Teknologi.
3. Bapak Eki Saputra, S.Kom., M.Kom sebagai Ketua Program Studi Sistem Informasi.
4. Ibu Siti Monalisa, ST., M.Kom sebagai Sekretaris Program Studi Sistem Informasi.
5. Bapak Tengku Khairil Ahsyar, S.Kom., M.Kom sebagai Kepala Laboratorium Program Studi Sistem Informasi.
6. Bapak Mustakim, ST., M.Kom sebagai Dosen Pembimbing Tugas Akhir ini.
7. Bapak Dr. Muhammad Luthfi Hamzah, B.IT., M.Kom sebagai Dosen Pembimbing Akademik peneliti yang telah banyak memberikan arahan, masukan, dan motivasi selama perkuliahan mulai dari Semester 1 hingga Semester 8 ini.
8. Ibu Fitriani Muttakin, S.Kom., M.Cs sebagai Penguji I dan Ibu Idria Maita, S.Kom., M.Sc sebagai Penguji II peneliti yang telah banyak memberikan arahan, masukan, nasihat, serta motivasi dalam penyelesaian Tugas Akhir ini juga dalam perkuliahan.
9. Seluruh Bapak dan Ibu Dosen Program Studi Sistem Informasi yang telah banyak memberikan ilmunya kepada peneliti. Semoga ilmu yang diberikan dapat peneliti amalkan dan menjadi amal jariyah.
10. Seluruh Pegawai dan Staf Fakultas Sains dan Teknologi yang telah membantu dan mempermudah proses administrasi selama perkuliahan ini.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Bapak Syahreza Al-Halsey sebagai Kepala Tata Usaha SMAN 1 Tualang yang telah meluangkan waktu dalam proses penelitian.
  2. Kedua orang tua peneliti yaitu Ayah Harjon dan Ibu Gusniarti tercinta yang tanpa lelah selalu memberikan semangat, motivasi, *support*, serta doa terbaiknya dan selalu menjadi motivasi peneliti dalam menyelesaikan Strata 1 (S1) ini.
  3. Abang Hardia Putra, Kakak Deva Achia, Kakak Ipar Nuraini, Abang Ipar Afrizal dan Keponakan Amara, Khaila, Alea, Luthfi, dan Ammar tercinta. Terima kasih telah memberikan perhatian, semangat, *support* serta doa kepada peneliti.
  4. Sahabat-sahabat tercinta peneliti yang namanya tak bisa disebutkan satu persatu.
  5. Keluarga besar Sistem Informasi 20 khususnya teman-teman dari SIF B Class, KKN Kelurahan Pangkalan Kerinci Timur 2023, serta teman seperbimbingan Tugas Akhir ini.
  16. Tim Riset dan Organisasi *Puzzle Research Data and Technology* (PREDATECH) yang banyak memberikan ilmu dan bimbingannya selama ini.
  17. Semua pihak yang namanya tidak dapat disebutkan satu persatu, yang telah banyak membantu dalam pelaksanaan serta penyelesaian Tugas Akhir ini.
- Semoga segala doa dan dorongan yang telah diberikan selama ini menjadi amal kebajikan dan mendapat balasan setimpal dari Allah *Subhanahu Wa Ta'ala*. Peneliti menyadari bahwa penulisan Tugas Akhir ini masih banyak terdapat kekurangan dan jauh dari kata sempurna. Untuk itu kritik dan saran yang membangun sangat diharapkan untuk kesempurnaan Tugas Akhir ini dan semoga Laporan Tugas Akhir ini bermanfaat bagi kita semua. Akhir kata peneliti ucapkan terima kasih.

*Wassalamu'alaikum Warahmatullahi Wabarakatuh*

Pekanbaru, 20 Mei 2024

Penulis,

UIN SUSKA RIAU

**NANDA TRY LUCHIA**

**NIM. 12050320445**

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

***FEATURE SELECTION PADA ALGORITMA SUPPORT VECTOR MACHINE DAN RANDOM FOREST UNTUK KLASIFIKASI PENERIMA PROGRAM INDONESIA PINTAR***

**NANDA TRY LUCHIA**  
**NIM: 12050320445**

Tanggal Sidang: 30 April 2024  
Periode Wisuda:

Program Studi Sistem Informasi  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Sultan Syarif Kasim Riau  
Jl. Soebrantas, No. 155, Pekanbaru

**ABSTRAK**

Pelaksanaan pendidikan bagi masyarakat Indonesia dari anak-anak hingga dewasa sangat dibutuhkan oleh para orang tua untuk mendidik dan membangun jiwa mereka. Banyaknya angka murid putus sekolah di Indonesia pada setiap tingkatan pendidikan menunjukkan bahwa masyarakat Indonesia masih memiliki tingkat pendidikan yang cukup rendah. Salah satu program yang dijalankan pemerintah adalah PIP yang juga terdapat di SMAN 1 Tualang. Penelitian ini dilakukan untuk mengklasifikasikan penerima bantuan pendidikan PIP menggunakan teknik klasifikasi menggunakan dua metode *Feature Selection* dan dua metode klasifikasi yaitu SVM dan RF. Hasil dari penelitian ini diperoleh bahwa penerapan *Feature Selection* pada Algoritma RF memiliki nilai akurasi yang lebih baik daripada Algoritma SVM dengan nilai akurasi tertinggi yaitu 99%, sedangkan Algoritma SVM memiliki nilai akurasi tertinggi yaitu 98%.

**Kata Kunci:** *Feature Selection*, Klasifikasi, PIP, *Random Forest*, SVM

# **FEATURE SELECTION IN SUPPORT VECTOR MACHINE AND RANDOM FOREST ALGORITHMS FOR THE CLASSIFICATION OF RECIPIENTS OF THE SMART INDONESIA PROGRAM**

**NANDA TRY LUCHIA**  
**NIM: 12050320445**

*Date of Final Exam: Apr 30<sup>th</sup> 2024*  
*Graduation Period:*

*Department of Information System*  
*Faculty of Science and Technology*  
*State Islamic University of Sultan Syarif Kasim Riau*  
*Soebrantas Street, No. 155, Pekanbaru*

## **ABSTRACT**

*The implementation of education for Indonesian people from children to adults is needed by parents to educate and build their souls. The number of dropouts in Indonesia at every level of education shows that Indonesian people still have a fairly low level of education. One of the programs run by the government is the Smart Indonesia Program (PIP) which is also available at SMAN 1 Tualang. This research was conducted to classify recipients of PIP education assistance using classification techniques using two Feature Selection methods and two classification methods namely SVM and Random Forest. The results of this study obtained that the application of Feature Selection in Random Forest algorithm has a better accuracy value than SVM algorithm with the highest accuracy value of 99%, while SVM algorithm has the highest accuracy value of 98%.*

**Keywords:** *Classification, Feature Selection, PIP, Random Forest, SVM*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## DAFTAR ISI

<b>LEMBAR PERSETUJUAN</b>	<b>ii</b>
<b>LEMBAR PENGESAHAN</b>	<b>iii</b>
<b>LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL</b>	<b>iv</b>
<b>LEMBAR PERNYATAAN</b>	<b>v</b>
<b>LEMBAR PERSEMBAHAN</b>	<b>vi</b>
<b>KATA PENGANTAR</b>	<b>vii</b>
<b>ABSTRAK</b>	<b>ix</b>
<b>ABSTRACT</b>	<b>x</b>
<b>DAFTAR ISI</b>	<b>xi</b>
<b>DAFTAR GAMBAR</b>	<b>xiv</b>
<b>DAFTAR TABEL</b>	<b>xv</b>
<b>DAFTAR SINGKATAN</b>	<b>xvi</b>
<b>PENDAHULUAN</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Perumusan Masalah.....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan .....	4
1.5 Manfaat.....	5
1.6 Sistematika Penulisan.....	5
<b>LANDASAN TEORI</b>	<b>6</b>
2.1 <i>Data Mining</i> .....	6
2.2 Pengelompokan <i>Data Mining</i> .....	6
2.3 Tahapan <i>Data Mining</i> .....	7
2.4 <i>Feature Selection</i> .....	7
2.5 Klasifikasi.....	7
2.6 Algoritma <i>Support Vector Machine</i> .....	8

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.7	Algoritma <i>Random Forest</i> .....	8
2.8	Program Indonesia Pintar (PIP) .....	9
2.9	Profil SMAN 1 Tualang .....	9
2.10	Visi SMAN 1 Tualang .....	11
2.11	Misi SMAN 1 Tualang .....	11
2.12	Tujuan Pendidikan SMAN 1 Tualang .....	12
2.13	Struktur Organisasi .....	12
2.14	Penelitian Terdahulu .....	14
	<b>METODOLOGI PENELITIAN</b> .....	<b>15</b>
3.1	Tahap Perencanaan .....	15
3.2	Tahap Pengumpulan Data .....	16
3.3	Tahap <i>Preprocessing</i> Data .....	16
3.4	Tahap Hasil dan Analisis .....	17
3.5	Tahap Dokumentasi .....	18
<b>4</b>	<b>ANALISIS DAN HASIL</b> .....	<b>19</b>
4.1	Analisis Pendahuluan .....	19
4.2	Pengumpulan Data .....	20
4.3	Penentuan Atribut .....	20
4.4	<i>Preprocessing</i> Data .....	20
	4.4.1 <i>Cleaning</i> Data .....	20
	4.4.2 Transformasi Data .....	21
	4.4.3 Normalisasi Data .....	22
4.5	Seleksi Fitur menggunakan <i>Chi Square</i> .....	22
4.6	Seleksi Fitur menggunakan <i>Information Gain Based</i> .....	23
4.7	Klasifikasi menggunakan SVM .....	24
	4.7.1 Klasifikasi SVM dengan <i>Feature Selection Chi Square</i> .....	24
	4.7.2 Klasifikasi SVM dengan <i>Feature Selection Information Gain Based</i> .....	25
	4.7.3 Klasifikasi SVM tanpa <i>Feature Selection</i> .....	26
4.8	Klasifikasi menggunakan RF .....	27
	4.8.1 Klasifikasi RF dengan <i>Feature Selection Chi Square</i> .....	27
	4.8.2 Klasifikasi RF dengan <i>Feature Selection Information Gain Based</i> .....	29
	4.8.3 Klasifikasi RF tanpa <i>Feature Selection</i> .....	30
4.9	Perbandingan Evaluasi .....	32
4.10	Diskusi Analisis .....	33

**Hak Cipta Dilindungi Undang-Undang**

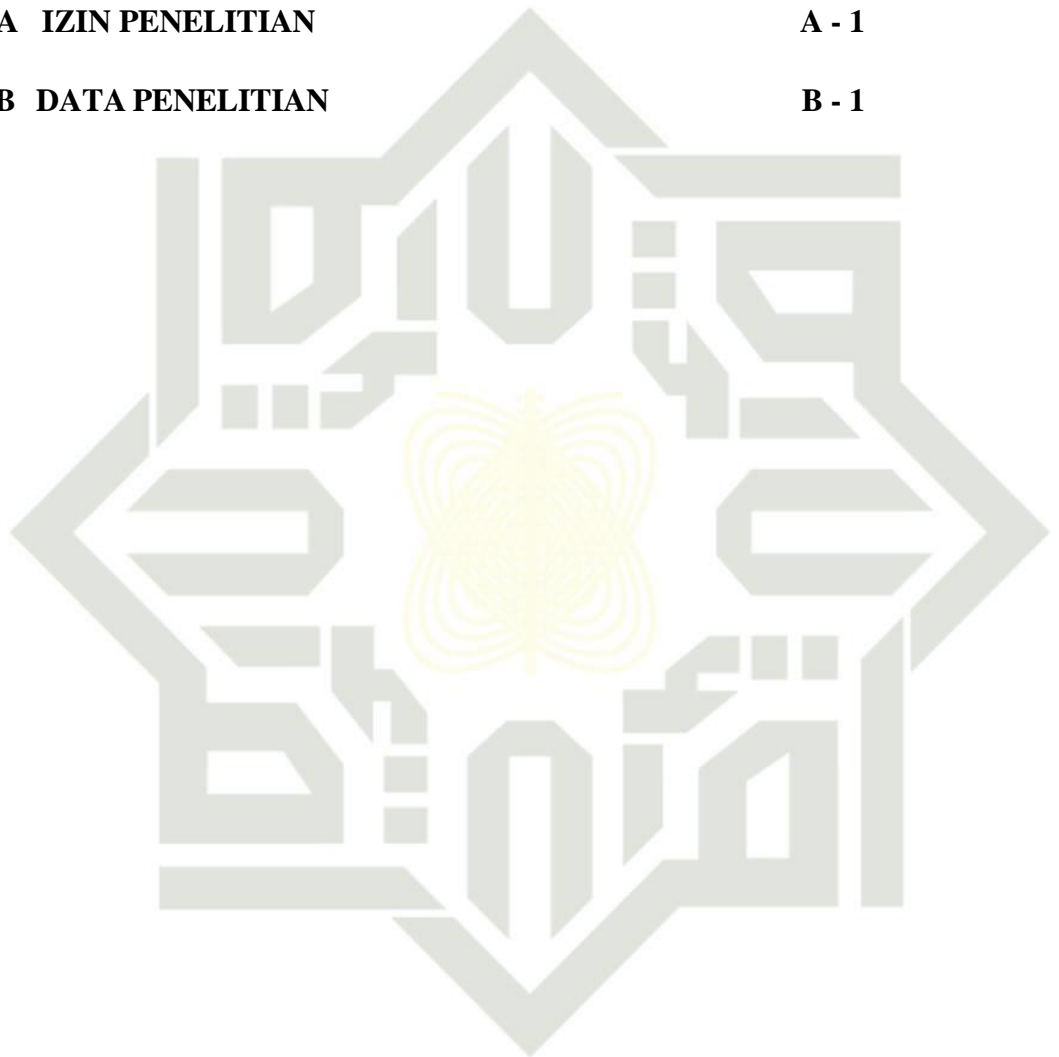
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4.11 Analisis Sistem Informasi . . . . .	35
<b>PENUTUP</b>	<b>37</b>
5.1 Kesimpulan . . . . .	37
5.2 Saran . . . . .	37

**DAFTAR PUSTAKA**

<b>LAMPIRAN A IZIN PENELITIAN</b>	<b>A - 1</b>
-----------------------------------	--------------

<b>LAMPIRAN B DATA PENELITIAN</b>	<b>B - 1</b>
-----------------------------------	--------------



## DAFTAR GAMBAR

1.1	Jumlah Peserta Didik Calon Penerima PIP Di SMAN 1 Tualang . . .	2
2.1	Struktur Organisasi SMAN 1 Tualang . . . . .	13
3.1	Metodologi Penelitian . . . . .	15
3.2	Tahapan Klasifikasi Algoritma SVM . . . . .	17
3.3	Tahapan Klasifikasi Algoritma rf . . . . .	17
4.1	Besar Penghasilan Orang Tua Murid SMAN 1 Tualang . . . . .	19
4.2	Syarat Pengujian SVM . . . . .	24
4.3	<i>Decision Tree</i> Dengan <i>Feature Selection Chi Square</i> . . . . .	28
4.4	<i>Decision Tree</i> Dengan <i>Feature Selection Information Gain Based</i> . . . . .	29
4.5	<i>Decision Tree</i> Tanpa <i>Feature Selection</i> . . . . .	31
4.6	Perbandingan Akurasi Klasifikasi Data <i>Kaggle</i> dan PIP . . . . .	33
4.7	Atribut Terpilih . . . . .	34
A.1	Surat Selesai Penelitian . . . . .	A - 1
A.2	Surat Keterangan Wawancara . . . . .	A - 2

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## DAFTAR TABEL

4.1	Hasil <i>Cleaning</i> Data . . . . .	21
4.2	Hasil Transformasi Data . . . . .	21
4.3	Hasil Normalisasi Data . . . . .	22
4.4	Hasil <i>Chi Square</i> . . . . .	23
4.5	Hasil <i>Information Gain Based</i> . . . . .	23
4.6	Hasil Akurasi SVM Dengan <i>Feature Selection Chi Square</i> . . . . .	25
4.7	Hasil Akurasi SVM Dengan <i>Feature Selection Information Gain Based</i> . . . . .	26
4.8	Hasil Akurasi SVM Tanpa <i>Feature Selection</i> . . . . .	27
4.9	Data Sampel Dengan <i>Feature Selection Chi Square</i> . . . . .	27
4.10	Hasil Akurasi RF Dengan <i>Feature Selection Chi Square</i> . . . . .	28
4.11	Data Sampel Dengan <i>Feature Selection Information Gain Based</i> . . . . .	29
4.12	Hasil Akurasi RF Dengan <i>Feature Selection Information Gain Based</i> . . . . .	30
4.13	Data Sampel Tanpa <i>Feature Selection</i> . . . . .	30
4.14	Hasil Akurasi RF Tanpa <i>Feature Selection</i> . . . . .	31
4.15	Hasil Akurasi Klasifikasi . . . . .	32
B.1	Data Murid SMAN 1 Tualang . . . . .	B - 1

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## DAFTAR SINGKATAN

BSM	: Bantuan Murid Miskin
Dapodik	: Data Pokok Pendidikan
FS	: <i>Feature Selection</i>
IDS	: <i>Intrusion Detection System</i>
IPTEK	: Ilmu Pengetahuan dan Teknologi
KDD	: <i>Knowledge Discovery in Database</i>
Kemendikbud	: Kementerian Pendidikan dan Kebudayaan
KIP	: Kartu Indonesia Pintar
KKS	: Kartu Keluarga Sehat
KLHK	: Kementerian Lingkungan Hidup dan Kehutanan
KM	: Kilometer
KPS	: Kartu Perlindungan Sosial
NBC	: <i>Naive Bayes Classifier</i>
NIK	: Nomor Induk Kependudukan
NISN	: Nomor Induk Siswa Nasional
PIP	: Program Indonesia Pintar
PNS	: Pegawai Negeri Sipil
Polri	: Kepolisian Negara Republik Indonesia
RF	: <i>Random Forest</i>
Rombel	: Rombongan Belajar
SDA	: Sumber Daya Alam
SDM	: Sumber Daya Manusia
SMA	: Sekolah Menengah Atas
SMAN	: Sekolah Menengah Atas Negeri
SMAS Pemda	: Sekolah Menengah Atas Swasta Pemerintah Daerah
SVM	: <i>Support Vector Machine</i>
TNI	: Tentara Nasional Indonesia
Waka	: Wakil Kepala Sekolah



## BAB 1

### PENDAHULUAN

#### 1. Latar Belakang

Pelaksanaan pendidikan bagi masyarakat Indonesia dari anak-anak hingga dewasa sangat dibutuhkan oleh para orang tua untuk mendidik dan membangun jiwa mereka baik lahir maupun batin (Sujana, 2019). Tahapan ini dimulai dari Taman Kanak-Kanak (TK), Sekolah Dasar (SD), Sekolah Menengah Pertama (SMP), Sekolah Menengah Atas (SMA) hingga duduk di bangku perkuliahan. Peran dasar yang sangat penting dari sebuah pendidikan yaitu menciptakan Sumber Daya Manusia (SDM) yang berkualitas dan membentuk budi pekerti serta *attitude* warga negara demi meningkatkan angka keberhasilan suatu negara yang dilihat dari tingkat pendidikan negara tersebut (Adikvika, Merlina, dan Mayangky, 2021). Dalam Undang-Undang Dasar tahun 1945 Bab XIII Pasal 31 ayat 1-5, menyatakan tentang hak pendidikan yang dimiliki oleh setiap warga Indonesia. Oleh karena itu, pemerintah memiliki kewajiban untuk memastikan setiap warga mendapatkan layanan pendidikan demi menjamin kualitas pendidikan yang bebas diskriminasi dan bermutu (Aliyyah, Ulfah, Herawati, Rachmadtullah, dan Asmara, 2020).

Berdasarkan situs Kementerian Pendidikan dan Kebudayaan (Kemendikbud), angka murid putus sekolah yang ada di Indonesia pada tahun 2020/2021 pada tingkat SD mencapai 44.516 murid, tingkat SMP sebanyak 11.378 murid, dan tingkat SMA sebanyak 13.879 murid. Banyaknya angka murid putus sekolah di Indonesia pada setiap tingkatan pendidikan menunjukkan bahwa masyarakat Indonesia masih memiliki tingkat pendidikan yang cukup rendah. Untuk mengatasi permasalahan tersebut, salah satu solusi yang dapat dilakukan oleh pemerintah adalah memberikan dukungan dana berupa Bantuan Pendidikan atau Beasiswa kepada murid yang layak mendapatkannya (Zulkarnain dan Murtani, 2021).

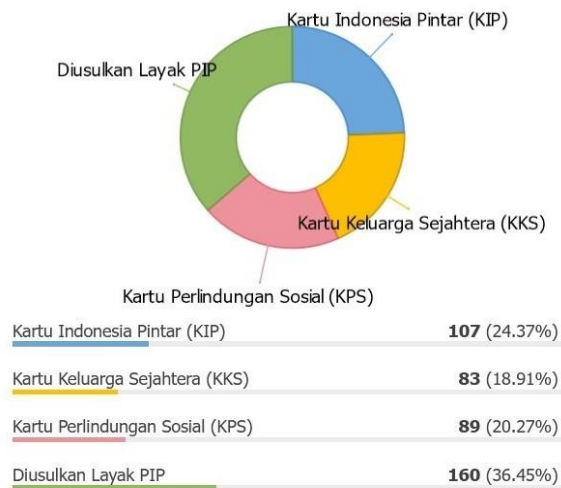
Bantuan pemerintah untuk para murid di daerah tertentu sudah dilakukan untuk membantu menunjang kebutuhan pendidikan yang berkebutuhan. Bantuan diberikan berdasarkan keadaan ekonomi dan juga prestasi yang dimiliki, baik dari segi akademik maupun di luar akademik (Cholil, Handayani, Prathivi, dan Ardanita, 2021). Bantuan ini bertujuan untuk membantu pihak terkait dalam keberlangsungan tingkat pendidikan yang ditempuh (Apriana, 2019). Hal tersebut dapat menjadi kesempatan bagi murid yang memiliki kendala ekonomi untuk tetap bisa memperoleh hak pendidikannya (Zaki, 2022). Salah satu program yang dijalankan pemerintah adalah Program Indonesia Pintar (PIP) yang juga terdapat di Sekolah Menengah Atas Negeri (SMAN) 1 Tualang.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**Hak Cipta Dilindungi Undang-Undang**

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

SMAN 1 Tualang merupakan sekolah dengan tingkat akreditasi A yang ditetapkan pada tanggal 02 November 2009. Sekolah ini sudah menjadi sekolah favorit di tingkat SMAN. Banyaknya murid berprestasi juga dihasilkan dari sekolah tersebut. PIP diberikan kepada murid yang memiliki kriteria miskin, konflik, kelainan fisik, keluarga terpidana, pemegang Kartu Indonesia Pintar (KIP), Kartu Perlindungan Sosial (KPS), Kartu Keluarga Sejahtera (KKS), hingga yatim piatu. Program ini berjalan setiap tahunnya dalam 4 tahap dan diberikan kepada murid yang layak. Dalam menentukan kelayakan penerima, pemegang KIP/KPS/KKS akan lebih diprioritaskan karena sudah tercatat di Data Pokok Pendidikan (Dapodik) sebagai keluarga kurang mampu. Sedangkan murid yang tidak memiliki ketiganya akan didata terlebih dahulu, dilakukan survei ke rumah masing-masing, dan diajukan usulan dari pihak sekolah sebagai murid yang layak mendapatkan bantuan. Grafik jumlah peserta didik calon penerima PIP di SMAN 1 Tualang dapat dilihat pada Gambar 1.1.



**Gambar 1.1.** Jumlah Peserta Didik Calon Penerima PIP Di SMAN 1 Tualang

Bantuan pendidikan dari pemerintah yang dilakukan selama ini, beberapa kali mengalami penyebaran yang kurang merata. Hal ini dapat disebabkan karena penerima bantuan kurang sesuai dengan kriteria yang seharusnya (Rizal dan Raharja, 2022). Permasalahan ini dapat menyebabkan penyebaran bantuan pendidikan untuk murid di suatu sekolah menjadi tidak merata. Untuk mengatasi masalah tersebut, penelitian ini dilakukan untuk mengklasifikasikan penerima bantuan pendidikan PIP menggunakan teknik klasifikasi. Teknik klasifikasi yang dapat digunakan untuk permasalahan tersebut yaitu *Support Vector Machine (SVM)* dan *Random Forest (RF)*.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

SVM merupakan teknik klasifikasi yang sering digunakan dalam beberapa tahun terakhir. SVM diperkenalkan sebagai Model *Machine Learning* berbasis *kernel* dengan pengenalan pola untuk melakukan klasifikasi dan *regresi* (Cervantes, Garcia-Lamont, Rodr'iguez-Mazahua, dan Lopez, 2020). Pada penelitian Sheykhmousa dan kawan-kawan, menganalisis SVM dan RF pada berbagai parameter. Hasil analisis menunjukkan bahwa SVM sangat dominan untuk melakukan klasifikasi (Sheykhmousa dkk., 2020). Menurut penelitian Ghosh dan kawan-kawan, SVM mengimplementasikan berbagai permasalahan dalam konsep yang luas. Hal ini menjadikan SVM sebagai teknik klasifikasi terbaik dalam penelitiannya (Ghosh, Dasgupta, dan Swetapadma, 2019). Teknik klasifikasi lainnya yaitu RF. RF merupakan Model *Machine Learning* yang dapat digunakan untuk melakukan klasifikasi dan prediksi (Speiser, Miller, Tooze, dan Ip, 2019). Dalam penelitian yang dilakukan Liu dan kawan-kawan menunjukkan bahwa RF tidak hanya mencapai target klasifikasi saja, melainkan kuantifikasi yang sangat efektif untuk kepentingan fitur yang ada (Liu dkk., 2021). Begitupula dengan penelitian yang dilakukan Amato dan kawan-kawan yang menemukan hasil klasifikasi menggunakan RF memberikan akurasi yang sempurna (Amato dkk., 2021).

Berdasarkan penelitian-penelitian yang telah dilakukan sebelumnya, menunjukkan bahwa Algoritma SVM dan RF memiliki keunggulan untuk melakukan klasifikasi. Oleh karena itu, penelitian ini dilakukan dengan membandingkan Algoritma SVM dan RF. Pada percobaan pertama, akan dilakukan klasifikasi dengan menggunakan data yang diperoleh dari *Kaggle*. Percobaan kedua dilakukan untuk mengklasifikasikan penerima PIP di SMAN 1 Tualang. Sebelum melakukan klasifikasi, penerapan *Feature Selection* dilakukan terlebih dahulu. Hal ini dilakukan untuk melihat apakah tiap atribut dapat mempengaruhi akurasi pada tiap algoritma (Puspadini, 2020). Dalam penelitian ini menggunakan Teknik *Chi Square* dan *Information Gain Based* untuk melakukan *Feature Selection* dari data yang diperoleh. *Feature Selection* dapat menghasilkan tingkat akurasi klasifikasi yang lebih baik (Sukarna dan Ansori, 2022). *Feature Selection* mampu menemukan atribut yang diprioritaskan untuk meningkatkan hasil analisis (Gurning, Permana, Maita, dkk., 2023). Penelitian ini diharapkan dapat menghasilkan analisis klasifikasi terbaik dari kedua percobaan data yang dimiliki. Sehingga, dapat diimplementasikan untuk meningkatkan kualitas pendidikan Indonesia dan membantu mengurangi angka putus sekolah di tingkat SMA.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## 1.2 Perumusan Masalah

Berdasarkan permasalahan diatas, didapatkan rumusan masalah yaitu bagaimana menerapkan *Chi Square* dan *Information Gain Based* untuk klasifikasi penerima bantuan PIP di SMAN 1 Tualang dengan menggunakan Algoritma SVM dan RF.

## 1.3 Batasan Masalah

Dalam melakukan analisis dan penelitian, perlu dituliskan batasan-batasannya agar masalah yang akan dibahas tidak menyimpang atau melahirkan masalah baru. Batasan masalah dalam penelitian ini adalah:

1. Studi kasus dalam penelitian ini adalah SMAN 1 Tualang, Kabupaten Siak, Provinsi Riau. Sumber data penelitian untuk percobaan pertama diperoleh dari *Kaggle* untuk melakukan *training* dan pengujian *performance* SVM dan RF, dan percobaan kedua diperoleh dari pihak sekolah yang bersangkutan. Jumlah data yang diperoleh sebanyak 1.112 murid, 26 atribut, dan 2 *class*. Atribut yang terdapat dalam penelitian ini yaitu Jenis Kelamin, Nomor Induk Siswa Nasional (NISN), Tempat Lahir, Tanggal Lahir, Nomor Induk Kependudukan (NIK), Agama, Alamat, Kelurahan, Kecamatan, Jenis Tinggal, Alat Transportasi, Penerima KPS, Pekerjaan Ayah, Penghasilan Ayah, Nama Ibu, Pekerjaan Ibu, Penghasilan Ibu, Rombongan Belajar (Rombel) Saat Ini, Penerima KIP, Layak PIP (usulan dari sekolah), Kebutuhan Khusus, Anak Keberapa, Berat Badan, Tinggi Badan, Jumlah Saudara Kandung, dan Jarak Rumah ke Sekolah (KM). Jenis data merupakan data terstruktur. Hasil penelitian membandingkan akurasi data *Kaggle* dan data SMAN 1 Tualang.
2. Data yang digunakan dalam penelitian adalah data seluruh murid SMAN 1 Tualang.
3. Proses pembagian data menggunakan Teknik *Hold Out* dengan percobaan 80:20, 70:30, dan 60:40.
4. Penelitian menggunakan dua teknik *Feature Selection*, yaitu *Chi Square* dan *Information Gain Based* untuk seleksi fitur data dan dua algoritma klasifikasi yaitu SVM dan RF.
5. Proses pengolahan data dan klasifikasi menggunakan *Python*.

## 1.4 Tujuan

Adapun tujuan dari Tugas Akhir ini adalah:

1. Menerapkan teknik *Chi Square* dan *Information Gain Based* untuk melihat kemampuan seleksi atribut data penerima bantuan PIP di SMAN 1 Tualang.

**Hak Cipta Dilindungi Undang-Undang**

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## **Manfaat**

Manfaat dari Tugas Akhir ini adalah:

1. Memberikan rekomendasi pola untuk menentukan penerima bantuan PIP di SMAN 1 Tualang.
2. Penelitian ini memberikan informasi dan manfaat di bidang rekomendasi beasiswa berdasarkan Metode SVM dan RF. Hasil penelitian ini kemudian dapat digunakan sebagai referensi tambahan untuk penelitian selanjutnya dan sebagai sumber informasi bagi yang berkepentingan dengan masalah tersebut.

## **Sistematika Penulisan**

Sistematika penulisan laporan ini terdiri dari lima bab yaitu:

### **BAB 1. PENDAHULUAN**

Bab ini merupakan pengantar laporan Tugas Akhir berupa latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan

### **BAB 2. LANDASAN TEORI**

Bab ini merupakan paparan teori yang diperoleh dari berbagai artikel, jurnal ilmiah, buku, serta studi kepustakaan sebagai rujukan dalam penulisan Tugas Akhir ini.

### **BAB 3. METODOLOGI PENELITIAN**

Bab ini memberikan gambaran alur yang digunakan dalam penelitian disertakan tahapan metode yang digunakan.

### **BAB 4. ANALISIS DAN HASIL**

Bab ini menjelaskan analisis data dari proses pengolahan data hingga mendapatkan hasil klasifikasi menggunakan Algoritma SVM dan RF.

### **BAB 5. PENUTUP**

Bab ini berisi kesimpulan Tugas Akhir dan saran penulis pada pembaca mengenai penelitian yang akan datang.



## BAB 2

### LANDASAN TEORI

#### 2.1 *Data Mining*

*Data Mining* merupakan proses yang digunakan dalam menemukan hubungan, pola, maupun prediksi dari data dengan jumlah besar yang dipilah dan memiliki nilai, diproses menggunakan teknologi pengenalan pola maupun statistik matematika (Yang dkk., 2020). *Data Mining* menghasilkan pengetahuan dari pemrosesan data di *database*. Istilah lain *Data Mining* yang dikenal memiliki makna yang sama yaitu *Knowledge Discovery in Database* (KDD) (Dogan dan Birant, 2021). *Data Mining* berguna untuk mengelola data dengan jumlah yang sangat besar demi membantu menyelesaikan pemrosesan data agar memperoleh informasi yang dibutuhkan dan diperlukan oleh pengguna (Rady dan Anwar, 2019). Penggunaan *Data Mining* sering dilakukan secara bergantian agar dapat menggali informasi yang sulit diperoleh dalam *database* yang besar (Ayatollahi, Gholamhosseini, dan Salehi, 2019).

#### 2.2 *Pengelompokan Data Mining*

*Data Mining* terbagi menjadi beberapa bagian (Yang dkk., 2020):

1. *Asosiasi*  
Asosiasi dalam *Data Mining* berguna untuk menemukan hubungan antar data dalam sejumlah data yang sangat besar dengan analisis.
2. *Clustering*  
Analisis *cluster* bekerja dengan mempelajari karakteristik yang serupa dalam sebuah data untuk di kelompokkan ke dalam satu kategori.
3. *Klasifikasi*  
Dalam *Data Mining*, klasifikasi bertujuan untuk mengekstrak data yang berharga dengan cara menandai data yang digunakan.
4. *Prediksi*  
Prediksi merupakan proses untuk mendapatkan hasil yang berguna di masa depan.
5. *Deskripsi*  
Pengolahan data yang dapat memberikan hasil berupa penjelasan pola yang muncul jika ada penambahan pada data disebut sebagai deskripsi.
6. *Estimasi*  
Metode estimasi menggunakan target proses yang lebih berfokus pada hasil numerik.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## 2.3 Tahapan *Data Mining*

*Data Mining* merupakan bagian dari proses KDD. Tahapan KDD adapun sebagai berikut (Sumantri dan Utami, 2020):

### 1. Pemilihan Data

Proses ini merupakan tahapan yang perlu dilakukan sebelum peneliti menggali informasi agar hasilnya dapat disimpan dalam *database* dan digunakan saat proses *Data Mining* berlangsung.

### 2. *Cleaning* Data

*Cleaning* data berfokus pada tindakan menyingkirkan data duplikat, data *noisy*, kesalahan pada data, dan lainnya agar informasi yang diperoleh lebih relevan dan bernilai.

### 3. Transformasi

Istilah *coding* dipakai saat proses transformasi dilakukan. Proses ini tergantung pada pola dan jenis informasi yang dibutuhkan oleh peneliti pada *database*.

### 4. *Data Mining*

Proses inti dalam KDD adalah *Data Mining*. Tahapan ini merupakan tahapan untuk menemukan pola dan informasi yang dibutuhkan dengan menggunakan metode, algoritma, maupun teknik tertentu pada sebuah data yang bervariasi.

### 5. Evaluasi

Informasi yang dihasilkan pada proses *Data Mining* disajikan dalam bentuk yang mudah dipahami.

## 2.4 *Feature Selection*

Seleksi atribut dilakukan agar pengolahan data dalam jumlah yang sangat besar berjalan cepat dan optimal karena *Data Mining* membutuhkan komputasi yang sangat tinggi (Khaire dan Dhanalakshmi, 2022). Fitur yang kurang relevan akan membuat proses klasifikasi tidak memberikan hasil yang maksimal karena fitur tersebut hanya bisa memberikan informasi yang sedikit bahkan tidak memiliki hubungan dengan data (Kou dkk., 2020). Dalam melakukan proses *Data Mining*, seleksi atribut akan membantu mengurangi nilai *error* untuk mempertahankan kualitas proses *Data Mining* agar hasil yang diperoleh tidak keluar dari jalur yang seharusnya (Raj dkk., 2020).

## 2.5 Klasifikasi

Teknik mendasar yang ada dalam *Data Mining* salah satunya adalah klasifikasi. Klasifikasi bisa menjadi solusi untuk berbagai kasus dalam membentuk suatu

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

aturan dalam sebuah data di dalam *database* (Ketui, Wisomka, dan Homjun, 2019). Pemecahan masalah dalam *Data Mining* menggunakan klasifikasi memiliki ketergantungan dengan data latih (Shafiq, Tian, Bashir, Jolfaei, dan Yu, 2020). Tujuan dilakukannya klasifikasi yaitu untuk memprediksi *class* suatu objek dari data yang ada di *database* (Umar Sidiq, Aaqib, dan Khan, 2019). Klasifikasi juga dapat memberikan gambaran untuk membedakan antara *class* data dan konsep berdasarkan pola analisis dari data latih yang diketahui (Pati, Parhi, dan Pattanayak, 2023). Strategi klasifikasi digunakan secara luas pada setiap bidang untuk mengklasifikasikan setiap data ke dalam *class* yang berbeda (Arifin dkk., 2021).

### 2.6 Algoritma Support Vector Machine

SVM merupakan teknik klasifikasi yang sering digunakan dalam beberapa tahun terakhir. SVM diperkenalkan sebagai Model *Machine Learning* berbasis *kernel* dengan pengenalan pola untuk melakukan klasifikasi dan regresi (Cervantes dkk., 2020). Metode SVM secara linier memisahkan *class* dengan menampilkan fitur ke dalam *kernel*. Secara umum, SVM bekerja dengan menambahkan ruang fitur dengan dimensi tambahan (Muhathir, Santoso, dan Larasati, 2021). SVM dapat menguji tingkat keakuratan data dalam sebuah sistem. Pembagian data yang dilakukan SVM dilakukan dengan cara mencari titik *vector* terdekat yang mendekati *hyperline* dari tiap kelompok (Tuhenay, 2021). Prinsip dasar pada Algoritma SVM yaitu klasifikasi *linear*. Hasil akurasi SVM bergantung pada parameter yang digunakan dan fungsi *kernel*nya (Amelia, Indra, dan Masruriyah, 2022). Dalam hal ini, persamaan *linear* SVM digunakan Persamaan 2.1.

$$Y = W^T xi + b, i = 1, 2, \dots, l \quad (2.1)$$

Dimana  $Y$  merupakan nilai target dari himpunan data berupa  $[-1, +1]$ ,  $xi$  adalah *vector* baris berdimensi  $k$  (jumlah fitur), sedangkan  $l$  adalah jumlah data,  $W$  merupakan parameter bobot, dan  $b$  yaitu bias atau *error* (Suryani dan Mustakim, 2022).

### 2.7 Algoritma Random Forest

RF merupakan Model *Machine Learning* yang dapat digunakan untuk melakukan klasifikasi dan prediksi (Speiser dkk., 2019). RF menggunakan Metode *Bagging* agar dapat memperbaiki proses prediksi yang dilakukan (Azhari, Situmorang, dan Rosnelly, 2021). Metode ini telah berhasil membuktikan penyelesaian masalah regresi dan klasifikasi selama beberapa tahun terakhir (Apriliah, Kurniawan, Baydhowi, dan Haryati, 2021). Hasil klasifikasi metode ini bergantung pada

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

pohon keputusan. RF terdiri dari *Root Node*, *Internal Node*, dan *Leaf Node* (Sandag, 2020). Pohon keputusan diperoleh sebagai penentu tingkat kemurnian atribut dengan menghitung nilai *entropy*. Persamaan yang digunakan untuk menghitung nilai *entropy* yaitu:

$$Entropy(Y) = - \sum_i p(c/Y) \log_2 p(c/Y) \quad (2.2)$$

Dimana  $Y$  merupakan kasus dan nilai  $Y$  pada *class c* (Siburian dan Mulyana, 2019).

### 2.8 Program Indonesia Pintar (PIP)

PIP merupakan bagian dari bantuan sosial yang menjadi kebijakan pemerintahan dalam bidang pendidikan untuk murid yang berasal dari kalangan miskin (Retnaningsih, 2019). Umumnya, PIP merupakan penyempurnaan dari program sebelumnya, yaitu Bantuan Murid Miskin (BSM) berupa bantuan uang tunai untuk pendidikan anak sekolah (Yusup, Ismanto, dan Wasitohadi, 2019). PIP diberikan kepada murid yang memiliki kriteria miskin, konflik, kelainan fisik, keluarga terpidana, pemegang KIP/KPS/KKS, hingga yatim piatu. Program ini berjalan setiap tahunnya dan diberikan kepada murid yang layak.

Dalam menentukan kelayakan penerima, pemegang KIP/KPS/KKS akan lebih diprioritaskan karena sudah tercatat di Dapodik sebagai keluarga kurang mampu. Sedangkan murid yang tidak memiliki ketiganya akan didata terlebih dahulu, dilakukan survei ke rumah masing-masing, dan diajukan usulan dari pihak sekolah sebagai murid yang layak mendapatkan bantuan. Dalam peraturan menteri, tujuan kebijakan ini dilakukan yaitu (Rakista, 2020):

1. Meningkatkan layanan pendidikan bagi anak usia sekolah untuk mendapatkan haknya dalam mengakses pendidikan wajib belajar selama 12 tahun.
2. Menurunkan angka murid putus sekolah akibat faktor ekonomi.
3. Memberikan kesempatan kedua bagi murid yang sudah putus sekolah untuk bisa melanjutkan jenjang pendidikannya.

### 2.9 Profil SMAN 1 Tualang

SMAN 1 Tualang merupakan sebuah sekolah menengah atas yang terletak di Kecamatan Tualang. Terbentuknya SMAN 1 Tualang berawal dari sebuah sekolah swasta yang bernama Sekolah Menengah Atas Swasta Pemerintah Daerah (SMAS Pemda) Perawang. Berdirinya sekolah swasta tersebut diprakarsai oleh masyarakat Tualang yang berasal dari masyarakat yang heterogen. Awal tujuan didirikannya SMAS Pemda Perawang yaitu untuk memenuhi kebutuhan masyarakat melanjutkan

#### Hak Cipta Dilindungi Undang-Undang

pendidikan tingkat perguruan tinggi.

Pada tahun 1997 SMAS Pemda Perawang merupakan satu-satunya sekolah SMA yang terdapat di Perawang. Selanjutnya, pada tahun 2002 SMAS Pemda Perawang ditetapkan oleh pemerintah beralih status dari sekolah swasta menjadi sekolah negeri dan berganti nama menjadi SMAN 1 Tualang. Kemudian, SMA ini berubah nama lagi menjadi SMAN 4 Siak pada tahun 2009 sampai tahun 2012. Tahun 2012, berubah kembali menjadi SMAN 1 Tualang hingga sekarang.

SMAN 1 Tualang merupakan sekolah favorit yang menjadi kebanggaan bagi masyarakat Tualang. Sekolah ini juga ditetapkan sebagai sekolah rujukan. SMAN 1 Tualang mendapatkan penghargaan dari Kementerian Lingkungan Hidup dan Kehutanan (KLHK) Republik Indonesia pada tahun 2018 sebagai Sekolah Adiwiyata tingkat nasional. SMAN 1 Tualang turut berkontribusi dalam upaya meningkatkan SDM di Kabupaten Siak khususnya, dan Provinsi Riau secara umumnya. Kontribusi tersebut terlihat dengan banyaknya peserta didik lulusan dari SMAN 1 Tualang yang diterima di perguruan-perguruan tinggi favorit yang tersebar di seluruh Indonesia.

Secara geografis, SMAN 1 Tualang terletak di Kecamatan Tualang yang merupakan pusat pertumbuhan ekonomi di Kabupaten Siak. SMAN 1 Tualang mudah dijangkau karena berada sekitar 50 kilometer (KM) dari pusat Pemerintah Provinsi Riau. Letaknya yang strategis membuat SMAN 1 Tualang didukung oleh akses komunikasi dan transportasi yang lancar. Tentunya kondisi ini juga berdampak positif terhadap perkembangan sekolah.

Secara demografis, peserta didik berasal dari berbagai macam bentuk kedudukan perekonomian karena orang tua peserta didik bermata pencarian heterogen seperti Pegawai Negeri Sipil (PNS), Tentara Nasional Indonesia (TNI), Kepolisian Negara Republik Indonesia (Polri), Karyawan, Anggota Dewan hingga pengusaha yang berkompeten di daerah setempat. Sebagian besar orang tua peserta didik bertempat tinggal di pusat keramaian dan pusat pertumbuhan ekonomi. Dari segi keterlibatan orang tua peserta didik dalam program-program sekolah, secara umum para orang tua peserta didik memiliki kepedulian tinggi terhadap sekolah.

Pada umumnya masyarakat memahami bahwa dalam rangka pengembangan sekolah sangat diperlukan adanya kontribusi dan kerja sama yang optimal antara sekolah dengan komite sekolah. Secara historis SMAN 1 Tualang memiliki sejarah yang cukup panjang karena SMAN 1 tualang menjadi barometer pendidikan tingkat SMA di Kabupaten Siak. Fenomena ini didukung oleh profesional dan keseriusan *staff* sekolah dalam mengembangkan sekolah. Dari prestasi yang diperoleh, sekolah ini menjadi sekolah terdepan dalam bidang prestasi akademik dan non-akademik di tingkat kabupaten. Pada tahun 2021 SMAN 1 Tualang menjadi

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

salah satu sekolah di Provinsi Riau yang masuk kategori 1.000 sekolah terbaik di Indonesia berdasarkan hasil nilai UTBK tahun 2021. SMAN 1 Tualang juga menjalin kerja sama dengan lembaga-lembaga pendidikan yang berkompeten sehingga mewujudkan visi dan misi SMAN 1 Tualang.

## 2.10 Visi SMAN 1 Tualang

Perkembangan dan tantangan masa depan seperti perkembangan ilmu pengetahuan dan teknologi, globalisasi yang sangat cepat, era informasi dan teknologi 4.0, serta perubahan tingkat kesadaran masyarakat dan orang tua terhadap dunia pendidikan memicu sekolah untuk merespon tantangan sekaligus peluang tersebut. SMAN 1 Tualang memiliki etika moral yang menggambarkan profil sekolah yang diinginkan di masa depan dengan visi sekolah yaitu “Mewujudkan SMAN 1 Tualang sebagai sekolah yang unggul, berwawasan global, berdaya saing, cinta lingkungan, dan berbudaya Melayu berlandaskan iman dan takwa”. Visi tersebut mencerminkan cita-cita sekolah yang berorientasi ke depan dengan memperhatikan potensi kekinian sesuai dengan norma dan harapan masyarakat. Untuk mewujudkannya, sekolah menentukan langkah-langkah strategis yang dinyatakan dalam misi sekolah.

## 2.11 Misi SMAN 1 Tualang

1. Mengembangkan sistem manajemen yang handal dan berkualitas.
2. Meningkatkan sistem pembelajaran yang berkualitas dan berhasil guna.
3. Meningkatkan kondisi lingkungan sekolah yang kondusif dan bernuansa religius.
4. Membentuk peserta didik yang beriman dan bertakwa.
5. Meningkatkan profesionalisme SDM yang beriman dan bertakwa.
6. Mengembangkan pembelajaran berbasis iman dan taqwa, Ilmu Pengetahuan dan Teknologi (IPTEK), dan akhlak mulia.
7. Mempersiapkan peserta didik untuk mengikuti pendidikan lebih lanjut, memiliki kecerdasan dan kompetensi untuk hidup mandiri, menguasai IPTEK, mampu bersaing di taraf regional, nasional, dan internasional.
8. Membentuk karakter peserta didik yang peduli terhadap lingkungan.
9. Meningkatkan kualitas lingkungan hidup, mencegah pencemaran, mengatasi kerusakan dan melakukan pelestarian terhadap Sumber Daya Alam (SDA) serta lingkungan.
10. Mewujudkan sekolah yang asri, aman, dan nyaman.

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## 2.12 Tujuan Pendidikan SMAN 1 Tualang

Pendidikan memiliki peran sentral bagi upaya pembangunan SDM. Oleh karena itu, penyelenggaraan pendidikan perlu dimutakhirkan sesuai dengan perkembangan dan kemajuan IPTEK. Implikasinya, jika ada pada saat ini masyarakat Indonesia dan dunia menghendaki tersedianya SDM yang memiliki seperangkat kompetensi berstandar Nasional dan Internasional. Maka isi dan proses pendidikannya perlu di arahkan pada pencapaian kompetensi tersebut.

Tujuan sekolah merupakan suatu keinginan yang hendak dicapai oleh suatu sekolah yang menjadi bagian dari tujuan pendidikan, guna meningkatkan kecerdasan, pengetahuan, kepribadian, budi pekerti dan akhlak mulia, serta keterampilan untuk hidup mandiri. Mempersiapkan peserta didik untuk dapat melanjutkan dan mengikuti pendidikan ke jenjang yang lebih tinggi hingga akhirnya dapat bersaing di era globalisasi. Peserta didik merupakan salah satu komponen penting bagi berlangsungnya kegiatan pendidikan di sekolah. Guru dan peserta didik merupakan komponen yang tidak dapat dipisahkan satu sama lainnya. Guru sebagai pendidik dan pengajar sedangkan siswa sebagai anak didik.

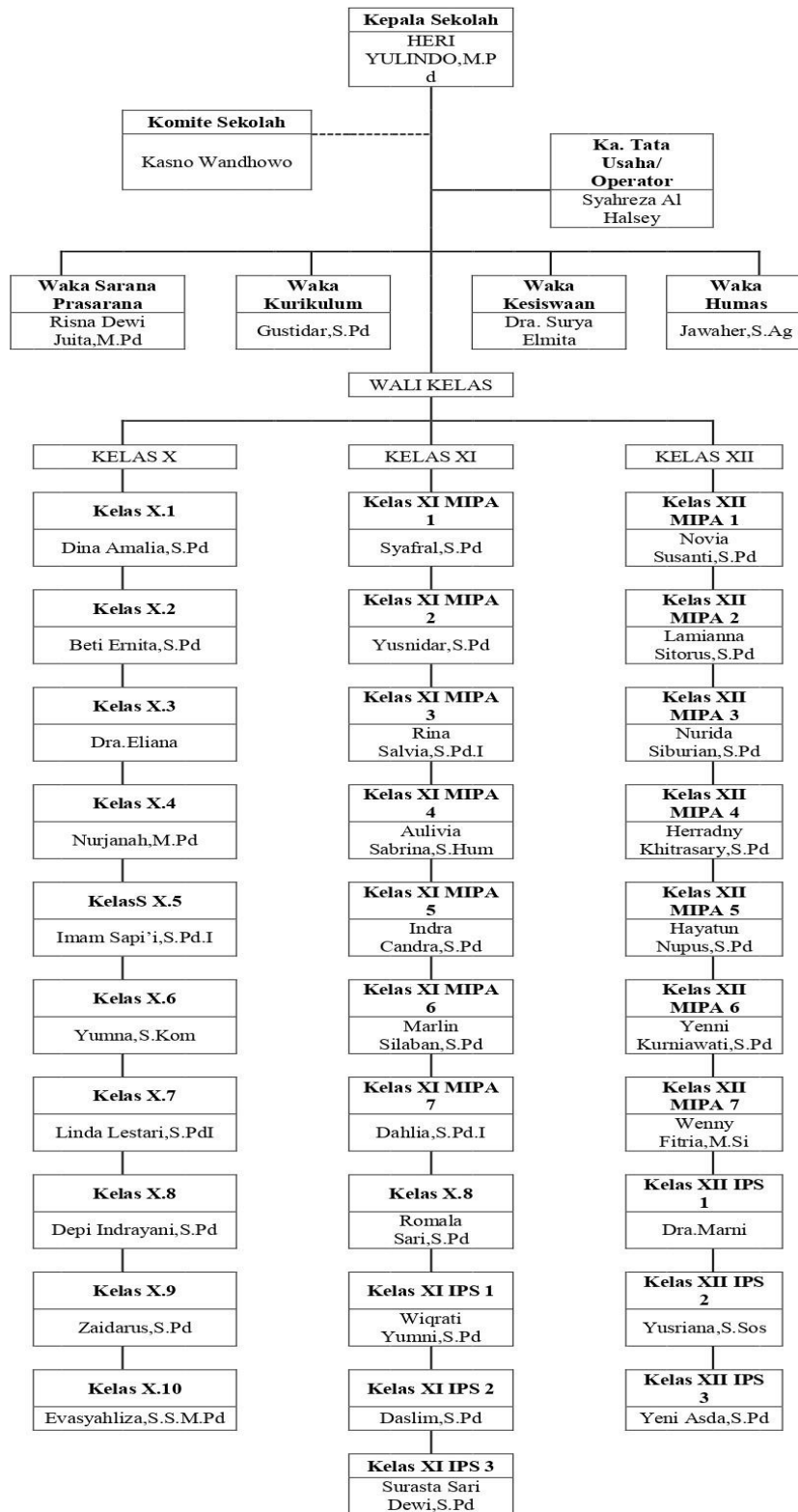
## 2.13 Struktur Organisasi

SMAN 1 Tualang di pimpin oleh Bapak Heri Yulindo, M.Pd selaku Kepala Sekolah dan Bapak Kasno Wandhowo sebagai Komite Sekolah. Terdapat 4 Wakil Kepala Sekolah (Waka) di bidang yang berbeda yaitu Ibu Risna Dewi Juita, M.Pd sebagai Waka Sarana Prasarana, Ibu Gustidar, S.Pd sebagai Waka Kurikulum, Ibu Dra. Surya Elmita sebagai Waka Kesiswaan, dan Ibu Jawaher, S.Ag sebagai Waka Humas. Terdapat 1.112 murid di SMAN 1 Tualang. Keseluruhan data peserta didik ini dipegang oleh Kepala Tata Usaha SMAN 1 Tualang, Bapak Syahreza Al Halsey. Dalam penelitian ini, peneliti melakukan wawancara langsung ke SMAN 1 Tualang dengan menemui Bapak Syahreza Al Halsey yang juga menjadi penanggung jawab PMP dan Administrator di SMAN 1 Tualang.

Bapak Syahreza Al Halsey memberikan perspektif yang sangat berharga berdasarkan pengalaman dan pengetahuannya yang luas di bidang tersebut. Beliau menjelaskan berbagai aspek dengan rinci, mulai dari permasalahan yang dihadapi hingga strategi efektif yang dapat diusulkan untuk mengatasinya. Informasi yang disampaikan oleh Bapak Syahreza Al Halsey sangat membantu dalam memperkaya pemahaman saya mengenai studi kasus dan permasalahan yang ada serta menambah nilai substansial pada penelitian yang saya lakukan. Keahlian dan keterbukaan beliau dalam berbagi pengetahuan menjadikan wawancara ini sangat produktif dan inspiratif. Struktur organisasi SMAN 1 Tualang dapat dilihat pada Gambar 2.1.

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



**Gambar 2.1.** Struktur Organisasi SMAN 1 Tualang

#### Hak Cipta Dilindungi Undang-Undang

## 2.14 Penelitian Terdahulu

Pada penelitian Sheykhmousa dan kawan-kawan, menganalisis SVM dan RF pada berbagai parameter. Hasil analisis menunjukkan bahwa SVM sangat dominan untuk melakukan klasifikasi (Sheykhmousa dkk., 2020). Menurut penelitian Ghosh dan kawan-kawan, SVM mengimplementasikan berbagai permasalahan dalam konsep yang luas. Hal ini menjadikan SVM sebagai teknik klasifikasi terbaik dalam penelitiannya (Ghosh dkk., 2019). Adrian dan kawan-kawan juga melakukan penelitian analisis sentimen dengan membandingkan algoritma SVM dan RF. Hasil klasifikasi diperoleh bahwa SVM mampu mendeteksi *label* “Positif” lebih baik daripada RF (Adrian, Putra, Rafialdy, dan Rakhmawati, 2021). Pada tahun 2019, Riadi dan kawan-kawan melakukan analisis perbandingan *Intrusion Detection System* (IDS) menggunakan SVM dan *Naïve Bayes*. Dalam penelitian ini diperoleh grafik dengan nilai akurasi terbaik ada pada SVM (Riadi, Umar, dan Aini, 2019).

Dalam penelitian yang dilakukan Liu dan kawan-kawan menunjukkan bahwa RF tidak hanya mencapai target klasifikasi saja, melainkan kuantifikasi yang sangat efektif untuk kepentingan fitur yang ada (Liu dkk., 2021). Begitupula dengan penelitian yang dilakukan Amato dan kawan-kawan yang menemukan hasil klasifikasi menggunakan RF memberikan akurasi yang sempurna (Amato dkk., 2021). Pada tahun 2021, Religia dan kawan-kawan melakukan analisis perbandingan menggunakan optimasi pada RF untuk mengklasifikasikan Data *Bank Marketing*. Dalam penelitian ini diperoleh akurasi performa paling optimal sebesar 88,30% (Religia, Nugroho, Hadikristanto, dkk., 2021). Irawan dan kawan-kawan melakukan perbandingan klasifikasi menggunakan 4 algoritma yaitu SVM, *Naive Bayes Classifier* (NBC), RF, dan *Bagging Classifier*. Hasil yang diperoleh dari perbandingan tersebut menunjukkan RF memiliki *Performance Score* yang cukup baik yaitu 96,8% (Irawan, Perkasa, Yurindra, Wahyuningsih, dan Helmud, 2021).

Shrawan Kumar Trivedi, seorang peneliti asal India melakukan sebuah penelitian pemodelan terkait penggunaan dan pengaruh seleksi fitur terhadap pendekatan *Machine Learning*. Penelitian ini dilakukan pada tahun 2020 dengan menggunakan tiga metode seleksi fitur yaitu *Information Gain*, *Gain Ratio*, dan *Chi Square*. Bukan hanya itu, ia menguji ketiga metode tersebut pada lima jenis algoritma klasifikasi seperti *Bayesian*, NBC, RF, *Decision Tree*, dan SVM. Kombinasi yang ia lakukan dalam penelitiannya memberikan bukti bahwa seleksi fitur dapat mempengaruhi keakuratan *Machine Learning* dalam melakukan klasifikasi (Trivedi, 2020).

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

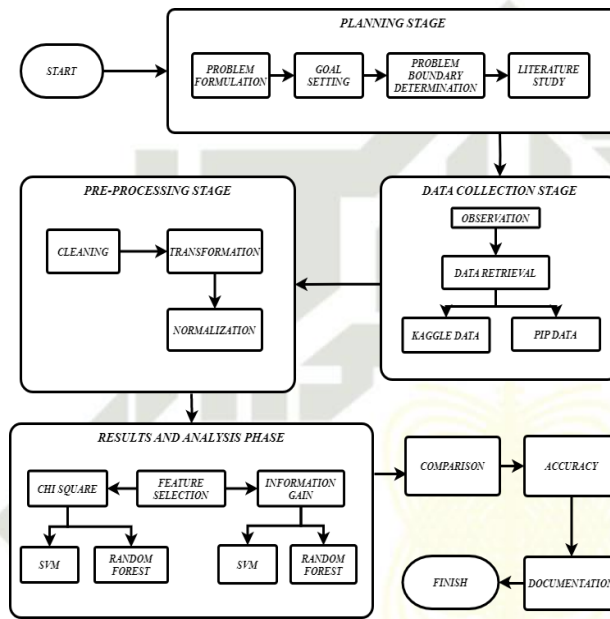
- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

### BAB 3

## METODOLOGI PENELITIAN

Penelitian ini terdiri dari 5 tahapan yaitu perencanaan, pengumpulan data, *preprocessing* data, hasil dan analisis, serta dokumentasi. Pemrosesan data terletak pada tahapan hasil dan analisis. Tahapan ini akan berfokus pada penerapan seleksi fitur pada tiap algoritma klasifikasi. Metodologi penelitian dalam penelitian ini dapat dilihat pada Gambar 3.1 dibawah ini.



Gambar 3.1. Metodologi Penelitian

#### 3.1 Tahap Perencanaan

Tahapan perencanaan merupakan tahapan yang dilakukan dalam sebuah penelitian untuk memperoleh perencanaan awal. Tahapannya terdiri dari:

1. Perumusan Masalah  
Rumusan masalah berisi masalah utama dari penelitian yang akan dilakukan yaitu menerapkan *Chi Square* dan *Information Gain Based* untuk klasifikasi penerima bantuan PIP di SMAN 1 Tualang dengan menggunakan Algoritma SVM dan RF.
2. Penentuan Tujuan  
Tujuan penelitian ditetapkan agar penelitian dapat menghasilkan sesuatu. Penelitian ini bertujuan untuk menerapkan Teknik *Chi Square* dan *Information Gain Based* untuk melihat kemampuan seleksi fitur pada data penerima bantuan PIP di SMAN 1 Tualang menggunakan Algoritma SVM dan RF.

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarangi mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3. Penentuan Batasan Masalah  
Batasan masalah bertujuan untuk membatasi isu atau masalah yang dihadapi agar tidak menghasilkan masalah baru.

4. Studi Literatur  
Studi literatur memberikan teori yang dipelajari dalam proses penelitian yang dapat diperoleh dari berbagai artikel, buku, jurnal, dan sumber lainnya.

### 3.2 Tahap Pengumpulan Data

Tahapan mengumpulkan data dilakukan untuk mendapatkan data atau informasi yang dibutuhkan sebagai bahan dalam penelitian. Metode yang dilakukan dalam pengumpulan data yaitu:

1. Observasi  
Observasi adalah metode pengumpulan data dengan cara terjun langsung ke lapangan dan mempelajari objek penelitian. Dalam penelitian ini dilakukan dengan cara mendatangi SMAN 1 Tualang secara langsung.
2. Pengambilan Data  
Metode ini dilakukan dengan dua cara. Pertama, mengambil data untuk *training* yang diperoleh dari *Kaggle*. Kedua, pengambilan data secara langsung ke SMAN 1 Tualang bersama narasumber yaitu data seluruh murid sebanyak 1.112 murid, 26 atribut, dan 2 *class*. Data yang digunakan dalam penelitian adalah data murid SMAN 1 Tualang saat ini.

### 3.3 Tahap Preprocessing Data

Tahapan *preprocessing* data menjadi tahap awal sebelum melakukan klasifikasi terhadap data-data yang telah dikumpulkan. Ada tiga hal yang harus dilakukan dalam *preprocessing* data yaitu:

1. *Cleaning* Data  
Proses *cleaning* dilakukan untuk memastikan data bersih, tidak ada data kosong, menghapus data yang tidak digunakan, dan mencegah terjadinya duplikat data.
2. Transformasi Data  
Transformasi data dilakukan untuk membuat data menjadi normal sehingga data dapat lebih optimal untuk di proses.
3. Normalisasi Data  
Proses normalisasi dilakukan agar tiap atribut dari pengolahan data memiliki rentang yang kecil dan sama. Pada penelitian ini, normalisasi yang dilakukan yaitu normalisasi *min-max*.

Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

### 3.4 Tahap Hasil dan Analisis

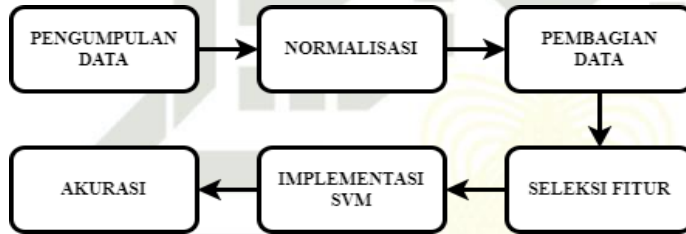
Setelah data mengalami *preprocessing*, maka perlu dilakukan pemrosesan data dalam penelitian. Beberapa hal yang dilakukan dalam tahapan ini yaitu:

1. *Feature Selection*

Proses ini dilakukan untuk menyeleksi fitur menggunakan dua metode yaitu *Chi Square* dan *Information Gain Based*. Sehingga bisa membandingkan tingkat akurasi yang akan diperoleh kedua algoritma berdasarkan dua percobaan data.

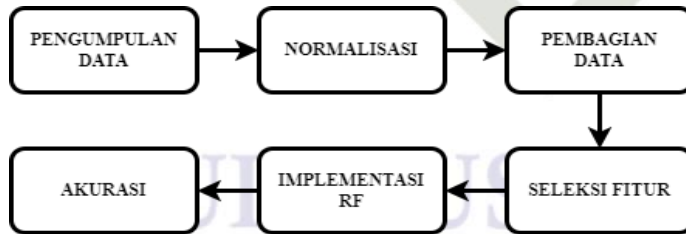
2. Klasifikasi SVM dan RF

Implementasi Algoritma SVM dan RF untuk mendapatkan hasil akurasi berdasarkan dua percobaan data yang telah diperoleh pada masing-masing metode seleksi fitur. Tahapan klasifikasi menggunakan SVM terdapat pada Gambar 3.2.



Gambar 3.2. Tahapan Klasifikasi Algoritma SVM

Gambar tersebut menunjukkan tahapan melakukan klasifikasi dengan Algoritma SVM yang dimulai dengan pengumpulan data, normalisasi, pembagian data, seleksi fitur, implementasi SVM, dan langkah terakhir yaitu mendapatkan nilai akurasi *performance* algoritma. Sedangkan tahapan melakukan klasifikasi menggunakan Algoritma RF terdapat pada Gambar 3.3.



Gambar 3.3. Tahapan Klasifikasi Algoritma rf

Gambar 3.3 menunjukkan tahapan melakukan klasifikasi dengan Algoritma RF yang dimulai dengan pengumpulan data, normalisasi, pembagian data, seleksi fitur, implementasi RF, dan terakhir akurasi.

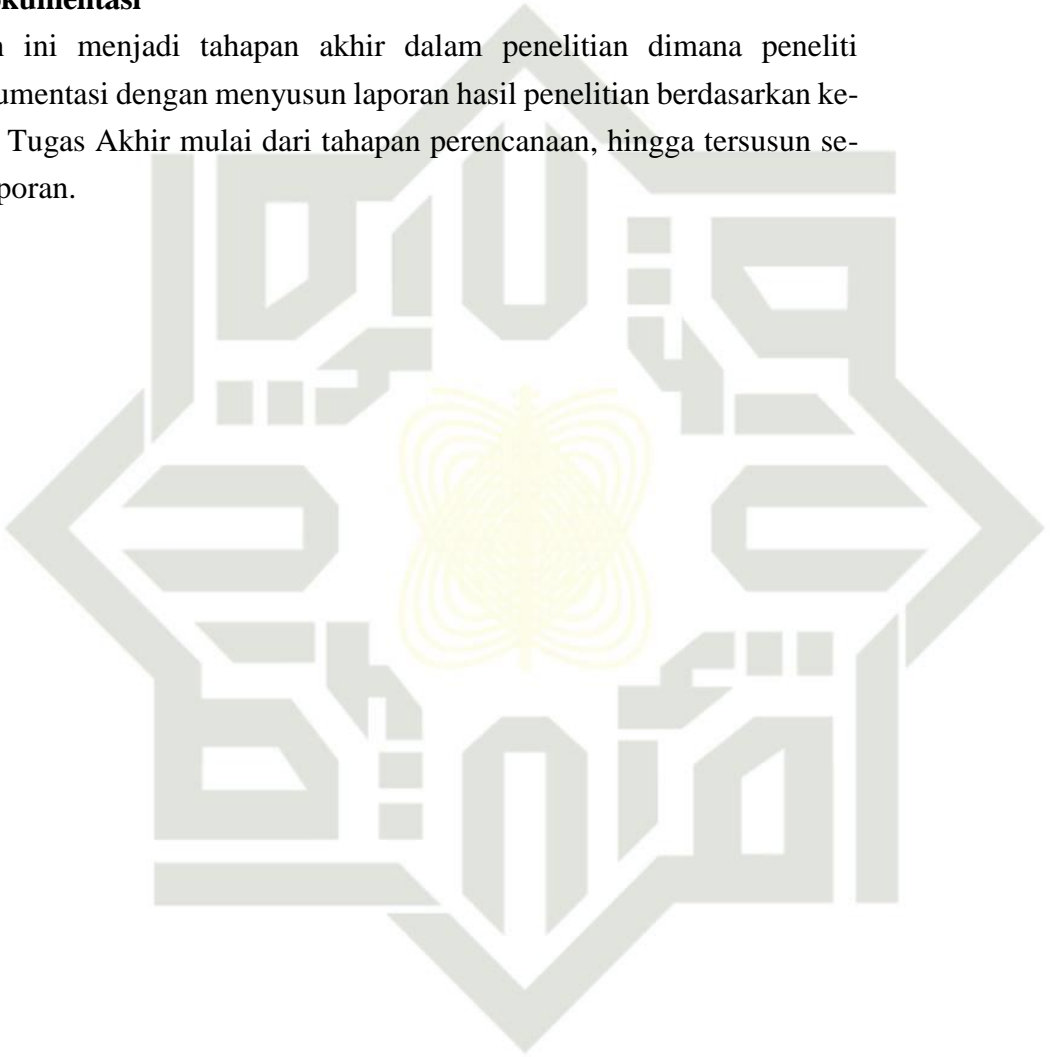
**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3. Perbandingan hasil klasifikasi kedua percobaan dengan Algoritma SVM dan RF.
4. Pada tahapan selanjutnya dilakukan perbandingan setiap akurasi yang telah dihasilkan dari Algoritma SVM dan RF dengan dua metode yaitu *Chi Square* dan *Information Gain Based*. Kedua data akan mendapatkan hasil akurasi yang berbeda.

**3. Tahap Dokumentasi**

Tahapan ini menjadi tahapan akhir dalam penelitian dimana peneliti melakukan dokumentasi dengan menyusun laporan hasil penelitian berdasarkan ketentuan laporan Tugas Akhir mulai dari tahapan perencanaan, hingga tersusun sebagai sebuah laporan.





## BAB 5 PENUTUP

### 5.1 Kesimpulan

Berdasarkan penelitian yang dilakukan untuk mengklasifikasikan data penerima PIP di SMAN 1 Tualang, diperoleh hasil bahwa Algoritma RF lebih baik dalam melakukan klasifikasi data penerima PIP di SMAN 1 Tualang dengan *Feature Selection Information Gain Based* dibandingkan dengan SVM. Hal ini dibuktikan dengan nilai rata-rata akurasi yang dihasilkan RF pada teknik *Feature Selection Information Gain Based* lebih besar daripada SVM yaitu 99%. Nilai ini merupakan nilai tertinggi di antara nilai rata-rata akurasi lainnya. Sehingga menunjukkan bahwa RF lebih baik dibandingkan SVM dalam klasifikasi data penerima PIP di SMAN 1 Tualang. Penelitian ini juga memiliki beberapa keterbatasan seperti data dikumpulkan melalui wawancara langsung ke satu studi kasus, penelitian selanjutnya dapat dilakukan dari sumber lain yaitu fokus dengan lebih banyak studi kasus. Hal ini membuka jalan bagi penelitian selanjutnya untuk melihat faktor-faktor lain yang dapat mempengaruhi rekomendasi penerimaan bantuan PIP di berbagai sekolah yang ada di Indonesia.

### 5.2 Saran

Saran penelitian ini untuk penelitian selanjutnya dapat dilakukan dengan melihat atribut-atribut yang digunakan di sekolah lain dalam penerimaan bantuan PIP dan menggunakan metode klasifikasi maupun prediksi untuk memperoleh hasil rekomendasi yang sempurna. Penambahan metode seleksi fitur juga dapat dilakukan seperti Metode *Fisher Score*, *Correlation Coefficient* dan *Variance Threshold*. Penelitian berikutnya diharapkan juga dapat membuat sistem pendukung keputusan berdasarkan metode yang telah digunakan dalam penelitian ini.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



## DAFTAR PUSTAKA

- Adikvika, A., Merlina, N., dan Mayangky, N. A. (2021). Sistem pendukung keputusan pemilihan penerima beasiswa pendidikan dengan menggunakan metode weighted product di yatim mandiri. *Indonesian Journal on Software Engineering (IJSE)*, 7(2), 148–158.
- Adrian, M. R., Putra, M. P., Rafialdy, M. H., dan Rakhmawati, N. A. (2021). Perbandingan metode klasifikasi random forest dan svm pada analisis sentimen psbb. *Jurnal Informatika Upgris*, 7(1).
- Ahyyah, R. R., Ulfah, S. W., Herawati, E. S. B., Rachmadtullah, R., dan Asmara, A. S. (2020). Bidikmisi: Analisis pelaksanaan program beasiswa pendidikan tinggi. *Journal Of Administration and Educational Management (ALIGNMENT)*, 3(1), 37–54.
- Amato, F., Coppolino, L., Cozzolino, G., Mazzeo, G., Moscato, F., dan Nardone, R. (2021). Enhancing random forest classification with nlp in dameh: A system for data management in ehealth domain. *Neurocomputing*, 444, 79–91.
- Amelia, U., Indra, J., dan Masruriyah, A. (2022). Implementasi algoritma support vector machine (svm) untuk prediksi penyakit stroke dengan atribut berpengaruh. *Scientific Student Journal for Information, Technology and Science*, 3(2), 254–259.
- Apriana, V. (2019). Penerapan profile matching untuk menentukan pemberian beasiswa pada siswa sekolah menengah atas. *Jurnal Mantik Penusa*, 3(1, Juni).
- Apriliah, W., Kurniawan, I., Baydhowi, M., dan Haryati, T. (2021). Prediksi kemungkinan diabetes pada tahap awal menggunakan algoritma klasifikasi random forest. *Sistemasi: Jurnal Sistem Informasi*, 10(1), 163–171.
- Arifin, M., Bhawika, G. W., Habibi, M. M., Firdaus, W., Agustinova, D. E., dan Rahim, R. (2021). Application of the cluster classification data mining method to child illiteracy in indonesia. *Library Philosophy and Practice*, 1–6.
- Atatollahi, H., Gholamhosseini, L., dan Salehi, M. (2019). Predicting coronary artery disease: a comparison between two data mining algorithms. *BMC public health*, 19(1), 1–9.
- Azhari, M., Situmorang, Z., dan Rosnelly, R. (2021). Perbandingan akurasi, recall, dan presisi klasifikasi pada algoritma c4. 5, random forest, svm dan naive bayes. *Jurnal Media Informatika Budidarma*, 5(2), 640–651.
- Cervantes, J., Garcia-Lamont, F., Rodr'iguez-Mazahua, L., dan Lopez, A. (2020).

Hak Cipta Dilindungi Undang-Undang

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim I

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.


**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, 408, 189–215.

Cholil, S. R., Handayani, T., Prathivi, R., dan Ardianita, T. (2021). Implementasi algoritma klasifikasi k-nearest neighbor (knn) untuk klasifikasi seleksi penerima beasiswa. *IJCIT (Indonesian Journal on Computer and Information Technology)*, 6(2), 118–127.

Dogan, A., dan Birant, D. (2021). Machine learning and data mining in manufacturing. *Expert Systems with Applications*, 166, 114060.

Chosh, S., Dasgupta, A., dan Swetapadma, A. (2019). A study on support vector machine based linear and non-linear pattern classification. Dalam *2019 international conference on intelligent sustainable systems (iciss)* (hal. 24–28).

Gurning, U. R., Permana, I., Maita, I., dkk. (2023). Chi-square features selection in unsupervised learning algorithm for measuring key performance indicators in riau province. Dalam *2023 international conference on computer science, information technology and engineering (iccosite)* (hal. 246–251).

Irawan, D., Perkasa, E. B., Yurindra, Y., Wahyuningsih, D., dan Helmud, E. (2021). Perbandingan klasifikasi sms berbasis support vector machine, naive bayes classifier, random forest dan bagging classifier. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 10(3), 432–437.

Ketui, N., Wisomka, W., dan Homjun, K. (2019). Using classification data mining techniques for students performance prediction. Dalam *2019 joint international conference on digital arts, media and technology with ecti northern section conference on electrical, electronics, computer and telecommunications engineering (ecti damt-ncon)* (hal. 359–363).

Khaire, U. M., dan Dhanalakshmi, R. (2022). Stability of feature selection algorithm: A review. *Journal of King Saud University-Computer and Information Sciences*, 34(4), 1060–1073.

Kou, G., Yang, P., Peng, Y., Xiao, F., Chen, Y., dan Alsaadi, F. E. (2020). Evaluation of feature selection methods for text classification with small datasets using multiple criteria decision-making methods. *Applied Soft Computing*, 86, 105836.

Lu, K., Hu, X., Zhou, H., Tong, L., Widanage, W. D., dan Marco, J. (2021). Feature analyses and modeling of lithium-ion battery manufacturing based on random forest classification. *IEEE/ASME Transactions on Mechatronics*, 26(6), 2944–2955.

Muhathir, M., Santoso, M. H., dan Larasati, D. A. (2021). Wayang image classification using svm method and glcm feature extraction. *Journal Of Informatics*



*And Telecommunication Engineering*, 4(2), 373–382.

- Pati, A., Parhi, M., dan Pattanayak, B. K. (2023). A review on prediction of diabetes using machine learning and data mining classification techniques. *International Journal of Biomedical Engineering and Technology*, 41(1), 83–109.
- Puspadini, R. (2020). Seleksi atribut pada algoritma k-nearest neighbor menggunakan similarity distance measures.
- Rady, E.-H. A., dan Anwar, A. S. (2019). Prediction of kidney disease stages using data mining algorithms. *Informatics in Medicine Unlocked*, 15, 100178.
- Raj, R. J. S., Shobana, S. J., Pustokhina, I. V., Pustokhin, D. A., Gupta, D., dan Shankar, K. (2020). Optimal feature selection-based medical image classification using deep learning model in internet of medical things. *IEEE Access*, 8, 58006–58017.
- Rakista, P. M. (2020). Implementasi kebijakan program indonesia pintar (pip). *Sawala: Jurnal Administrasi Negara*, 8(2), 224–232.
- Religia, Y., Nugroho, A., Hadikristanto, W., dkk. (2021). Analisis perbandingan algoritma optimasi pada random forest untuk klasifikasi data bank marketing. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(1), 187–192.
- Retnaningsih, H. (2019). Program indonesia pintar: Implementasi kebijakan jaminan sosial bidang pendidikan. *Aspirasi: Jurnal Masalah-masalah Sosial*, 8(2), 161–177.
- Riadi, I., Umar, R., dan Aini, F. D. (2019). Analisis perbandingan detection traffic anomaly dengan metode naive bayes dan support vector machine (svm). *ILKOM Jurnal Ilmiah*, 11(1), 17–24.
- Rizal, A., dan Raharja, W. K. (2022). Pemilihan pegawai penerima beasiswa pendidikan menggunakan algoritma topsis dan iso 9126 pada badan nasional penanggulangan bencana (bnpb). (*JurTI) Jurnal Teknologi Informasi*, 6(1), 64–70.
- Sandag, G. A. (2020). Prediksi rating aplikasi app store menggunakan algoritma random forest. *CogITo Smart Journal*, 6(2), 167–178.
- Safiq, M., Tian, Z., Bashir, A. K., Jolfaei, A., dan Yu, X. (2020). Data mining and machine learning methods for sustainable smart cities traffic classification: A survey. *Sustainable Cities and Society*, 60, 102177.
- Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanes, F., Ghamisi, P., dan Homayouni, S. (2020). Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 6308–6325.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



## Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- Siburian, V. W., dan Mulyana, I. E. (2019). Prediksi harga ponsel menggunakan metode random forest. Dalam *Annual research seminar (ars)* (Vol. 4, hal. 144–147).
- Speiser, J. L., Miller, M. E., Tooze, J., dan Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. *Expert systems with applications*, *134*, 93–101.
- Srijana, I. W. C. (2019). Fungsi dan tujuan pendidikan indonesia. *Adi Widya: Jurnal Pendidikan Dasar*, *4*(1), 29–39.
- Sukarna, R. H., dan Ansori, Y. (2022). Implementasi data mining menggunakan metode naive bayes dengan feature selection untuk prediksi kelulusan mahasiswa tepat waktu. *Jurnal Ilmiah Sains dan Teknologi*, *6*(1), 50–61.
- Sumantri, R. B. B., dan Utami, E. (2020). Penentuan status tahapan keluarga sejahtera kecamatan sidareja menggunakan teknik data mining. *Respati*, *15*(3), 71–82.
- Sryani, S., dan Mustakim, M. (2022). Estimasi keberhasilan siswa dalam pemodelan data berbasis learning menggunakan algoritma support vector machine. *Bulletin of Informatics and Data Science*, *1*(2), 81–88.
- Trivedi, S. K. (2020). A study on credit scoring modeling with different feature selection and machine learning approaches. *Technology in Society*, *63*, 101413.
- Tuhenay, D. (2021). Perbandingan klasifikasi bahasa menggunakan metode naive bayes classifier (nbc) dan support vector machine (svm). *JIKO (Jurnal Informatika dan Komputer)*, *4*(2), 105–111.
- Umar Sidiq, D., Aaqib, S. M., dan Khan, R. A. (2019). Diagnosis of various thyroid ailments using data mining classification techniques. *Int J Sci Res Coput Sci Inf Technol*, *5*, 131–6.
- Yang, J., Li, Y., Liu, Q., Li, L., Feng, A., Wang, T., ... Lyu, J. (2020). Brief introduction of medical database and data mining technology in big data era. *Journal of Evidence-Based Medicine*, *13*(1), 57–69.
- Yusup, W. B., Ismanto, B., dan Wasitohadi, W. (2019). Evaluasi program indonesia pintar dalam peningkatan akses pendidikan di sekolah menengah pertama. *Kelola: Jurnal Manajemen Pendidikan*, *6*(1), 44–53.
- Zaki, S. A. (2022). Analisis kerjasama sge-disdik jawa barat dengan jepang dalam program beasiswa pendidikan.
- Zulkarnain, W., dan Murtani, A. (2021). Analisis implementasi penyaluran dana zis untuk beasiswa pendidikan (studi kasus: Lazismu medan). *JURNAL AL-QASD ISLAMIC ECONOMIC ALTERNATIVE*, *2*(1), 11–20.



- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
    - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
    - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
  2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LAMPIRAN A IZIN PENELITIAN



**PEMERINTAH PROVINSI RIAU  
DINAS PENDIDIKAN  
SEKOLAH MENENGAH ATAS ( SMA ) NEGERI 1 TUALANG**  
 Jl. Sultan Alamuddinyah KM. 7 Perawang, Kabupaten Siak, Propinsi Riau Kode Pos 28772  
 Telp. : 081275434241 Website: www.sman1tualang.sch.id E-mail: smansatualang@gmail.com  
 NSS : 301091104006 NPSN : 10403425

### SURAT KETERANGAN

Nomor : 422/421.3- SMAN.1 Tualang/III/2024/346

Yang bertanda tangan dibawah ini Kepala SMA Negeri 1 Tualang Kecamatan Tualang Kabupaten Siak Provinsi Riau :

Nama : H. Heri Yulindo, M.Pd.  
 NIP : 19720316 200701 2 007  
 Pangkat/Golongan : Pembina Tk. I/IV. b  
 Jabatan : Kepala SMA Negeri 1 Tualang

Berdasarkan surat izin dari Universitas Islam Negeri Sultan Syarif Kasim Riau tanggal 15 Mei 2023, dengan ini kami SMA Negeri 1 Tualang memberi izin pengadakan penelitian pada Semester Genap 2023/2024 kepada mahasiswa berikut :

Nama : NANDA TRY LUCHIA  
 NIM : 12050320445  
 No Hp : 085274971412  
 Tempat/Tgl Lahir : Petapahan/08 Mei 2002  
 Prodi/Semester : Sistem Informasi/6  
 Instansi : Sekolah Menengah Atas Negeri (SMAN) 1 Tualang  
 Alamat Instansi : Jl. Sultan Alamuddinyah KM 7  
 Bidang/Judul : PENERAPAN FEATURE SELECTION DALAM PENENTUAN PENERIMA PROGRAM INDONESIA PINTAR (PIP) MENGGUNAKAN ALGORITMA NAÏVE BAYES DENGAN TEKNIK KLASIFIKASI

Untuk mendapatkan data yang berhubungan dengan kegiatan penelitian sampai dengan kegiatan tersebut selesai dilaksanakan.

Demikianlah surat keterangan ini dibuat agar digunakan sebagaimana mestinya.

Tualang, 06 Maret 2024  
 Kepala SMA Negeri 1 Tualang  
  
 H. Heri Yulindo, M.Pd  
 NIP. 19720316 200701 2 007

**Gambar A.1.** Surat Selesai Penelitian



### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## SURAT KETERANGAN TELAH MELAKUKAN WAWANCARA PENGAMBILAN DATA

Yang bertanda tangan dibawah ini :

Nama : SYAHREZA AL HALSEY  
NIP : 19781221 200701 1 002  
Jabatan : Ka. Tata Usaha

Menerangkan bahwa mahasiswa yang beridentitas dibawah ini :

Nama : NANDA TRY LUCHIA  
NIM : 12050320445  
No Hp : 085274971412  
Tempat/Tgl Lahir : Petapahan/08 Mei 2002  
Alamat : Perum Kopkar Jalan 2 No 17, Kec Tualang, Kab.Siak, Riau  
Email UIN : [nandaluchia@gmail.com](mailto:nandaluchia@gmail.com)

Benar telah melakukan wawancara pengambilan data, dalam rangka penelitian Tugas Akhir dengan judul " PENERAPAN FEATURE SELECTION DALAM PENENTUAN PENERIMA PROGRAM INDONESIA PINTAR (PIP) MENGGUNAKAN ALGORITMA NAÏVE BAYES DENGAN TEKNIK KLASIFIKASI ".

Demikianlah surat keterangan ini dibuat agar digunakan sebagaimana mestinya.

Tualang, 6 Maret 2024  
Ka. Tata Usaha SMAN. 1 Tualang

SYAHREZA AL HALSEY  
NIP : 19781221 200701 1 002

**Gambar A.2.** Surat Keterangan Wawancara



## LAMPIRAN B DATA PENELITIAN

**Tabel B.1.** Data Murid SMAN 1 Tualang

NO	NAMA	GENDER	NISN	...	CLASS
1	Aan Setiawan	L	0053593654	...	Layak
2	Abdi Anugrah Pratama	L	0063535740	...	Tidak Layak
3	Abdi Rizqullah Widodo	L	0046233481	...	Tidak Layak
4	Abdul Aziz	L	0067094505	...	Tidak Layak
5	Abdul Fattah	L	3052069457	...	Tidak Layak
6	Abdul Hamid	L	0056188566	...	Layak
7	Abdul Rahman	L	0057431347	...	Layak
8	Abdullah Abiyoso	L	0045832334	...	Tidak Layak
9	Adam Arsyad Chaniago	L	0046097762	...	Tidak Layak
10	Adam Somahua. Mb	L	0067300212	...	Tidak Layak
11	Ade Irma	P	0069807646	...	Layak
12	Ade Meutya Pinasih	P	0054552581	...	Tidak Layak
14	Adelia Putri Chalisa	P	0062429038	...	Tidak Layak
15	Adelya Indah Rahmah Ningsih	P	0053826221	...	Tidak Layak
16	Adhitya Dzahabi Elfman	L	0066611734	...	Tidak Layak
17	Adib Aldiansyah	L	0056605428	...	Tidak Layak
18	Adib Pratama	L	0069981672	...	Tidak Layak
19	Aditya Aghaci. E	L	0066995454	...	Tidak Layak
20	Aditya Pratama	L	0058375623	...	Layak
...	...	...	...	...	...
1110	Yunita Sari	P	0058779632	...	Layak
1111	Zhafira Rahadatul Aishi	P	0067076572	...	Layak
1112	Zuyyin Najwa	P	0046233431	...	Layak

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim I

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



## DAFTAR RIWAYAT HIDUP



© Ha

IN Suska Riau

State Islamic University of Sultan Syarif Kasim I

Nanda Try Luchia lahir di Desa Petapahan, pada tanggal 08 Mei 2002. Peneliti merupakan anak dari Bapak Harjon dan Ibu Gusniarti. Peneliti adalah anak ketiga dari tiga bersaudara. Pada tahun 2007 peneliti memulai pendidikan dengan masuk TK Bina Karya di Desa Perawang Barat dan lulus pada 20 Juni 2008. Lalu melanjutkan pendidikan Sekolah Dasar di SD Negeri 03 Perawang Barat. Peneliti menyelesaikan pendidikan Sekolah Dasar pada 21 Juni 2014. Setelah menyelesaikan pendidikan Sekolah Dasar peneliti melanjutkan pendidikan tingkat SLTP di SMP Negeri 1 Tualang yang selesai pada 02 Juni 2017. Peneliti melanjutkan pelajaran ke tingkat SLTA di SMAN 1 Tualang. Setelah menyelesaikan pendidikan di SMAN 1 Tualang pada 5 Mei 2020, peneliti pun melanjutkan pendidikan dengan menjadi mahasiswa Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Pada tahun 2022, peneliti telah menyelesaikan kerja praktek yang bertempat di PT Arara Abadi Kecamatan Tualang Kabupaten Siak, kemudian selanjutnya pada tahun 2023 peneliti juga telah menyelesaikan Kuliah Kerja Nyata (KKN) di Kelurahan Pangkalan Kerinci Timur, Kecamatan Pangkalan Kerinci, Kabupaten Pelalawan. Dan peneliti telah menyelesaikan kuliah Strata satu (S1) dengan mengambil judul tugas akhir “*Feature Selection Pada Algoritma Support Vector Machine Dan Random Forest Untuk Klasifikasi Penerima Program Indonesia Pintar*” pada tahun 2024.

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.