

**Analisis Pemilihan Jurusan pada Calon Siswa SMK
Negeri 4 Palembang Berdasarkan *Association Rule* pada
Faktor-Faktor Penentu Pemilihan Jurusan Menggunakan
*Random Forest***



**OLEH :
MUHAMAD HAFIZ BUDI PRAYOGA
09012682327021**

**MAGISTER ILMU KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2025**

**Analisis Pemilihan Jurusan pada Calon Siswa SMK
Negeri 4 Palembang Berdasarkan *Association Rule* pada
Faktor-Faktor Penentu Pemilihan Jurusan Menggunakan
*Random Forest***

TESIS

**Diajukan Untuk Melengkapi Salah Satu Syarat Memperoleh Gelar
Magister**



OLEH :

MUHAMAD HAFIZ BUDI PRAYOGA

09012682327021

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2025

HALAMAN PENGESAHAN

**Analisis Pemilihan Jurusan pada Calon Siswa SMK Negeri 4
Palembang Berdasarkan *Association Rule* pada Faktor-Faktor
Penentu Pemilihan Jurusan Menggunakan *Random Forest***

TESIS

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Magister

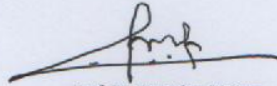
OLEH :

MUHAMAD HAFIZ BUDI PRAYOGA

09012682327021

Palembang, 9 Januari 2025

Pembimbing



Prof. Dr. Ermatita, M.Kom

NIP 196709132006042001

Mengetahui,

Koordinator Program Studi Magister Ilmu Komputer



Dr. Sirdaus, M.Kom

NIP 197804212008121003

Halaman Persetujuan

Pada hari Kamis, tanggal 02 Januari 2025 telah dilaksanakan ujian sidang tesis oleh Magister Ilmu Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Muhamad Hafiz Budi Prayoga
NIM : 09012682327021
Judul : Analisis Pemilihan Jurusan pada Calon Siswa SMK Negeri 4 Palembang Berdasarkan *Association Rule* pada Faktor-Faktor Penentu Pemilihan Jurusan Menggunakan *Random Forest*

1. Ketua Penguji

Dr. Firdaus, M.Kom

NIP 197801212008121003

2. Penguji I

Dr. Ali Ibrahim, S.Kom., M.T

NIP 198407212019031004

3. Penguji II

Dr. Abdiansah, S.Kom., M.CS

NIP 198410012009121005

4. Pembimbing

Prof. Dr. Ermatita, M.Kom

NIP 196709132006042001

Mengetahui,
Koordinator Program Studi Magister Ilmu Komputer



Dr. Firdaus, M.Kom

NIP 197801212008121003

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Muhamad Hafiz Budi Prayoga
NIM : 09012682327021
Program Studi : Magister Ilmu Komputer
Judul Tesis : Analisis Prediksi Jurusan pada siswa SMK Negeri 4 Palembang
Berdasarkan Association Rule pada Faktor-Faktor Penentu
Pemilihan Jurusan Menggunakan Metode Random Forest
Hasil Pengecekan Software iThenticate/Turnitin : 13 %

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

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, 20 Desember 2024



Muhamad Hafiz Budi Prayoga
NIM. 09012682327021

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala limpahan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan proposal penelitian ini yang berjudul **"Analisis Pemilihan Jurusan pada Calon Siswa SMK Negeri 4 Palembang Berdasarkan *Association Rule* pada Faktor-Faktor Penentu Pemilihan Jurusan Menggunakan *Random Forest*".**

Proposal ini tidak akan tersusun dengan baik tanpa dukungan dan bantuan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. Erwin, S.Si., M.Si selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
2. Bapak Dr. Firdaus, M.Kom selaku Kepala Program Studi Magister Ilmu Komputer
3. Bapak Hadi purnawan Satria, S.Kom, M.Sc, Ph.D Selaku Dosen Pembimbing Akademik
4. Ibu Prof. Dr. Ermatita, M.Kom selaku Dosen Pembimbing Tesis
5. Istri Tercinta Yellin Agustine, M.Pd.,Gr, Orang Tua, Adik, Saudara, dan Keluarga yang memberikan motivasi dan dukungan dalam menyelesaikan Pendidikan Magister Teknik Infomartika
6. Semua pihak yang tidak dapat disebutkan satu per satu, namun telah membantu baik secara langsung maupun tidak langsung dalam penyusunan penelitian ini.

Penulis menyadari bahwa dalam penyusunan proposal ini masih terdapat kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan dan dapat bermanfaat bagi pembaca serta dapat menjadi kontribusi positif bagi pengembangan keilmuan, khususnya di bidang pendidikan kejuruan.

Palembang, Januari 2025

Muhamad Hafiz Budi Prayoga

Analisis Pemilihan Jurusan pada Calon Siswa SMK Negeri 4 Palembang Berdasarkan *Association Rule* pada Faktor-Faktor Penentu Pemilihan Jurusan Menggunakan *Random Forest*

Muhamad Hafiz Budi Prayoga (09012682327021)

Master's Department of Computer Science, Faculty of Computer Science, Sriwijaya University

Email: 09012682327021@student.unsri.ac.id

Abstract

Choosing a major is a crucial decision for prospective students at SMK Negeri 4 Palembang, influenced by various academic and non-academic factors. This study aims to analyze the determinants of major selection using the Association Rule method and the Random Forest algorithm. Association Rule was utilized to identify patterns and relationships between key variables such as academic performance (Mathematics, Science, English, and Indonesian), while Random Forest was applied to validate and enhance the accuracy of the prediction model. The dataset used in this study consisted of 662 entries of prospective students from the 2024/2025 academic year, including academic scores, major preferences, and other relevant information. The analysis involved data preprocessing, the application of Association Rule to uncover significant patterns, and the use of Random Forest to develop a predictive model for major selection. The results showed that combining these two methods produced a high-accuracy prediction model and revealed strong correlations with factors influencing the decision-making process. This study provides strategic insights for schools to assist prospective students in selecting majors that align with their interests and potential. It also supports data-driven decision-making to improve the quality of vocational education.

Keywords: *Major Selection, SMK Negeri 4 Palembang, Association Rule, Random Forest, Data Mining*

Analisis Pemilihan Jurusan pada Calon Siswa SMK Negeri 4 Palembang Berdasarkan *Association Rule* pada Faktor-Faktor Penentu Pemilihan Jurusan Menggunakan *Random Forest*

Muhamad Hafiz Budi Prayoga (09012682327021)

Jurusan Magister Ilmu Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

Email: 09012682327021@student.unsri.ac.id

Abstrak

Pemilihan jurusan merupakan keputusan penting bagi calon siswa di SMK Negeri 4 Palembang, yang dipengaruhi oleh berbagai faktor akademik dan non-akademik. Penelitian ini bertujuan untuk menganalisis faktor-faktor penentu pemilihan jurusan menggunakan metode *Association Rule* dan algoritma *Random Forest*. *Association Rule* digunakan untuk mengidentifikasi pola dan hubungan antara variabel-variabel penentu seperti nilai akademik (Matematika, IPA, Bahasa Inggris, dan Bahasa Indonesia), sementara *Random Forest* diterapkan untuk memvalidasi dan meningkatkan akurasi model prediksi. Dataset yang digunakan dalam penelitian ini berasal dari data calon siswa tahun ajaran 2024/2025, yang mencakup 662 entri dengan atribut nilai akademik, pilihan jurusan, serta informasi lainnya. Proses analisis meliputi tahap pra-pemrosesan data, penerapan *Association Rule* untuk menemukan pola-pola utama, dan penerapan *Random Forest* untuk membangun model prediksi pemilihan jurusan. Hasil penelitian menunjukkan bahwa kombinasi kedua metode ini menghasilkan model prediksi dengan akurasi tinggi dan relevansi kuat terhadap faktor-faktor yang memengaruhi keputusan pemilihan jurusan. Penelitian ini memberikan kontribusi dalam bentuk panduan strategis bagi pihak sekolah untuk membantu calon siswa dalam memilih jurusan yang sesuai dengan minat dan potensi mereka, serta mendukung pengambilan keputusan yang berbasis data untuk meningkatkan kualitas pendidikan kejuruan.

Kata Kunci: Pemilihan Jurusan, SMK Negeri 4 Palembang, *Association Rule*, *Random Forest*, Data Mining

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PERNYATAAN.....	iv
KATA PENGANTAR.....	vi
<i>Abstrak</i>	vi
Abstrak	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	xii
DAFTAR TABLE	xiv
Daftar Lampiran.....	xv
BAB I	1
PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Tujuan dan Manfaat.....	3
1.3.1 Tujuan Penelitian.....	3
1.3.2 Manfaat Penelitian	4
1.4 Batasan Masalah.....	4
1.5 Sistematika Penulisan.....	5
BAB II	Error! Bookmark not defined.
TINJAUAN PUSTAKA	Error! Bookmark not defined.
2.1 SMK NEGERI 4 Palembang.....	Error! Bookmark not defined.
2.2 Machine Learning.....	Error! Bookmark not defined.

2.3	Association Rule.....	Error! Bookmark not defined.
2.4	Random Forest.....	Error! Bookmark not defined.
2.5	Cara Kerja Random Forest	Error! Bookmark not defined.
2.6	Klasifikasi.....	Error! Bookmark not defined.
2.7	Normalisasi Data	Error! Bookmark not defined.
2.8	Confusion Matrix.....	Error! Bookmark not defined.
2.9	Penelitian Terdahulu	Error! Bookmark not defined.
BAB III.....		Error! Bookmark not defined.
METODOLOGI PENELITIAN		Error! Bookmark not defined.
3.1	Metode Penelitian.....	Error! Bookmark not defined.
3.2	Studi Literatur.....	Error! Bookmark not defined.
3.3	Pengumpulan Data.....	Error! Bookmark not defined.
3.4	Pra – Pemrosesan Data	Error! Bookmark not defined.
3.4.1	Pengecekan Nilai Hilang.....	Error! Bookmark not defined.
3.4.2	Penghapusan Data Duplikat	Error! Bookmark not defined.
3.4.3	Penghapusan Kolom Yang Tidak Digunakan.....	Error! Bookmark not defined.
3.4.4	Encoding Data.....	Error! Bookmark not defined.
3.4.5	Pemisahan Data.....	Error! Bookmark not defined.
3.5	Penerapan Association Rule	Error! Bookmark not defined.
3.5.1	Penentuan Asscoation Rule	Error! Bookmark not defined.
3.5.2	Perhitungan Association Rule	Error! Bookmark not defined.
3.5.3	Proses Penentuan Association Rule Berdasarkan 8 Jurusan	Error! Bookmark not defined.
3.6	Penerapan Random Forest.....	Error! Bookmark not defined.

3.6.1	Perhitungan Random Forest.....	Error! Bookmark not de
3.6.2	Perhitungan Akurasi.....	Error! Bookmark not defined.
3.7	Hasil dan Kesimpulan.....	Error! Bookmark not defined.
BAB IVError! Bookmark not defined.
HASIL DAN PEMBAHASAN	ixError! Bookmark not defined.
4.1	Pra – Pemrosesan Data Python.....	Error! Bookmark not defined.
4.1.1	Pengecekan Nilai Hilang.....	Error! Bookmark not defined.
4.1.2	Penghapusan Data Duplikat	Error! Bookmark not defined.
4.1.3	Penghapusan Kolom Yang Tidak Digunakan	Error! Bookmark not defined.
4.1.4	Encoding Data.....	Error! Bookmark not defined.
4.1.5	Pembagian Dataset	Error! Bookmark not defined.
4.2	Implementasi Association Rule	Error! Bookmark not defined.
4.2.1	Association Rule TKJ	Error! Bookmark not defined.
4.2.2	Association Rule TKR	Error! Bookmark not defined.
4.2.3	Association Rule TBSM	Error! Bookmark not defined.
4.2.4	Association Rule TP.....	Error! Bookmark not defined.
4.2.5	Association Rule TAV	Error! Bookmark not defined.
4.2.6	Association Rule TITL.....	Error! Bookmark not defined.
4.2.7	Association Rule DPIB	Error! Bookmark not defined.
4.2.8	Association Rule RPL	Error! Bookmark not defined.
4.3	Implementasi Random Forest.....	Error! Bookmark not defined.
4.3.1	Random Forest TKJ	Error! Bookmark not defined.
4.3.2	Random Forest TKR	Error! Bookmark not defined.
4.3.3	Random Forest TBSM	Error! Bookmark not defined.

4.3.4	Random Forest TP.....	Error! Bookmark not defi
4.3.5	Random Forest TAV.....	Error! Bookmark not defi
4.3.6	Random Forest TITL.....	Error! Bookmark not defined.
4.3.7	Random Forest DPIB	Error! Bookmark not defined.
4.3.8	Random Forest RI _x	Error! Bookmark not defined.
4.4	Hasil dan Pembahasan.....	Error! Bookmark not defined.
BAB V.....		Error! Bookmark not defined.
KESIMPULAN DAN SARAN		Error! Bookmark not defined.
5.1	Kesimpulan.....	Error! Bookmark not defined.
5.2	Saran.....	Error! Bookmark not defined.
DAFTAR PUSTAKA.....		18

DAFTAR GAMBAR

HALAMAN

Gambar 3. 1	Diagram Alir Penelitian	Error! Bookmark not defined.
Gambar 3. 2	Diagram Alir Pra – Pemrosesan Data	Error! Bookmark not defined.
Gambar 4.1	Pengecekan Nilai Hilang	Error! Bookmark not defined.
Gambar 4. 2	Hasil Pengecekan Nilai Hilang	Error! Bookmark not defined.
Gambar 4.3	Penghapusan Data Duplikat.....	Error! Bookmark not defined.
Gambar 4.4	Hasil Penghapusan Kolom Yang Tidak Digunakan	Error! Bookmark not defined.
Gambar 4.5	Hasil Encoding Data	Error! Bookmark not defined.
Gambar 4.6	Mapping Kode Jurusan	Error! Bookmark not defined.
Gambar 4.7	Proses dan Hasil Association Rule TKJ	Error! Bookmark not defined.
Gambar 4.8	Proses dan Hasil Association Rule TKR	Error! Bookmark not defined.
Gambar 4.9	Proses dan Hasil Association Rule TSM	Error! Bookmark not defined.
Gambar 4.10	Proses dan Hasil Association Rule TP	Error! Bookmark not defined.
Gambar 4.11	Proses dan Hasil Association Rule TAV	Error! Bookmark not defined.
Gambar 4.12	Proses dan Hasil Association Rule TITLE	Error! Bookmark not defined.
Gambar 4.13	Proses dan Hasil Association Rule DPIB	Error! Bookmark not defined.
Gambar 4.14	Proses dan Hasil Association Rule RPL	Error! Bookmark not defined.
Gambar 4.15	Proses Untuk Random Forest Jurusan TKJ	Error! Bookmark not defined.

Gambar 4.16 Confusion Matrix TKJ**Error! Bookmark not defined.**

Gambar 4.17 Proses Untuk Random Forest Jurusan TKR**Error! Bookmark not defined.**

Gambar 4.18 Confusion Matrix TKR.....**Error! Bookmark not defined.**

Gambar 4. 19 Proses Untuk Random Forest Jurusan TBSM**Error! Bookmark not defined.**

Gambar 4. 20 Confusion Matrix TBSM.....**Error! Bookmark not defined.**

Gambar 4. 21 Proses Untuk Random Forest Jurusan TPE**Error! Bookmark not defined.**

Gambar 4. 22 Confusion Matrix TP**Error! Bookmark not defined.**

Gambar 4. 23 Proses Untuk Random Forest Jurusan TAVE**Error! Bookmark not defined.**

Gambar 4. 24 Confusion Matrix TAV**Error! Bookmark not de**

Gambar 4. 25 Proses Untuk Random Forest Jurusan TITL**Error! Bookmar defined.**

Gambar 4. 26 Confusion Matrix TITL**Error! Bookmark not defined.**

Gambar 4. 27 Proses Untuk Random Forest Jurusan DPIB**Error! Bookmark not defined.**

xii

Gambar 4. 28 Confusion Matrix D**Error! Bookmark not defined.**

Gambar 4. 29 Proses Untuk Random Forest Jurusan RPL**Error! Bookmark not defined.**

Gambar 4. 30 Confusion Matrix RPL**Error! Bookmark not defined.**

DAFTAR TABLE

HALAMAN

Table 2.1 Penelitian Terdahulu.....	Error! Bookmark not defined.
Table 3.1 Data Siswa SMK Negeri 4 Palembang ..	Error! Bookmark not defined.
Table 3.2 Hasil Pengecek Nilai Hilang	Error! Bookmark not defined.
Table 3. 3 Pengecek Data Duplikat	22
Table 3. 4 Penghapus Kolom Yang Tidak Digunakan	Error! Bookmark not defined.
Table 3. 5 Hasil Encoding Data.....	Error! Bookmark not defined.
Table 3. 6 Nilai Association Rule.....	Error! Bookmark not defined.
Table 3.7 Nilai Decision Tree TKR.....	Error! Bookmark not defined.
Table 3.8 Nilai Decision Tree TKJ.....	Error! Bookmark not defined.
Table 3.9 Nilai Decision Tree.....	Error! Bookmark not defined.
Table 3.10 Nilai Decision Tree.....	Error! Bookmark not defined.
Table 4.1 Ringkasan Pembagian Dataset	Error! Bookmark not defined.
Table 4.2 Association Rule Setiap Jurusan	Error! Bookmark not defined.
Table 4.3 Hasil Rekapitulasi Association Rule Jurusan	Error! Bookmark not defined.
Table 4.4 Hasil Rekapitulasi Random Forest	Error! Bookmark not defined.

DAFTAR LAMPIRAN

LAMPIRAN 1. Dataset

LAMPIRAN 2. Code

LAMPIRAN 3. Publikasi Ilmiah

LAMPIRAN 4. Hasil Pengecekan Software Ithenticate/Turnitin

LAMPIRAN 5. Surat Rekomendasi Ujian Tesis

LAMPIRAN 6. Form Revisi Ujian Tesis

LAMPIRAN 7. Form Konsultasi Tesis

LAMPIRAN 8. SK Persetujuan Pembimbing Tesis

LAMPIRAN 9. SK Pengangkatan Tim Penguji Tesis

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sekolah Menengah Kejuruan (SMK) adalah institusi pendidikan formal yang menyediakan pendidikan kejuruan di jenjang menengah. Tujuan dari pendidikan ini adalah memberikan keterampilan dan pengetahuan praktis kepada siswa agar mereka siap memasuki dunia kerja di bidang tertentu setelah menyelesaikan pendidikan (Setiawansyah, dkk 2020). SMK merupakan lanjutan dari Sekolah Menengah Pertama (SMP) atau Madrasah Tsanawiyah (MTs), atau bentuk pendidikan lain yang setara. Pendidikan di SMK lebih difokuskan pada pengembangan kompetensi praktis yang sesuai dengan kebutuhan industri dan pasar kerja, sehingga menghasilkan lulusan yang memiliki keterampilan spesifik dan siap kerja di berbagai bidang seperti teknik, bisnis, pariwisata, kesehatan, dan teknologi informasi.

SMK Negeri 4 Palembang adalah salah satu SMK di Kota Palembang yang menawarkan delapan jurusan yaitu Teknik Komputer Jaringan, Teknik Sepeda Motor, Teknik Kendaraan Ringan, Desain Pemodelan dan Informasi Bangunan, Teknik Instalasi Tenaga Listrik, Rekayasa Perangkat Lunak, Teknik Audio Visual, Teknik Permesinan. Beragamnya pilihan jurusan di SMK Negeri 4 Palembang memberikan banyak opsi bagi calon siswa. Pemilihan jurusan merupakan keputusan penting yang dapat mempengaruhi masa depan calon siswa. Di SMK Negeri 4 Palembang, berbagai jurusan disediakan untuk memenuhi kebutuhan industri dan minat calon siswa. Namun, keputusan ini sering dipengaruhi oleh berbagai faktor, baik internal maupun eksternal, seperti minat pribadi, prestasi akademik, dukungan keluarga, informasi dari teman sebaya, serta prospek kerja di masa depan.

Mengidentifikasi faktor-faktor yang paling berpengaruh dalam pemilihan jurusan dapat memberikan wawasan berharga bagi pihak sekolah untuk membantu calon siswa membuat keputusan yang tepat. Salah satu pendekatan untuk analisis ini adalah Association Rule, yang dapat mengidentifikasi hubungan atau asosiasi

antara berbagai faktor penentu. Selain itu, metode Random Forest dapat digunakan untuk memvalidasi dan meningkatkan akurasi prediksi pemilihan jurusan.

Association rules adalah metode yang digunakan untuk mengelompokkan data dalam jumlah besar, seperti kumpulan item, dengan cara memasukkan data tersebut ke dalam serangkaian aturan tertentu. Aturan-aturan ini kemudian digunakan oleh sistem untuk mengoreksi jawaban dari peserta didik. Dalam prosesnya, data yang terkumpul diurai dan dianalisis untuk menemukan pola-pola yang signifikan (Prasetyo, dkk 2022). Pola-pola ini kemudian dirumuskan menjadi aturan-aturan yang membantu sistem dalam menilai dan memberikan umpan balik yang akurat kepada peserta didik. Dengan demikian, association rules tidak hanya membantu dalam pengelompokan data tetapi juga dalam meningkatkan kualitas evaluasi dan pembelajaran.

Random forest adalah metode klasifikasi yang dikembangkan dari metode Decision Tree, di mana pemilihan atribut pada setiap node dilakukan secara acak untuk menentukan klasifikasi. Dalam proses klasifikasinya, keputusan akhir didasarkan pada mayoritas suara dari semua pohon keputusan yang telah dibentuk. Pembentukan random forest menggunakan teknik yang dikenal sebagai bagging, dengan pemilihan atribut secara acak. Untuk menumbuhkan pohon keputusan, digunakan metode CART (Classification and Regression Tree). Pohon-pohon ini akan tumbuh hingga ukuran maksimum dan tidak akan dipangkas, menghasilkan kumpulan pohon yang disebut sebagai hutan (forest) (Ratnawati dan Sulistyaningrum 2020).

Berdasarkan permasalahan ini, penulis memiliki gagasan untuk melakukan penelitian tentang "Analisis Pemilihan Jurusan Pada Calon Siswa SMK Negeri 4 Palembang Berdasarkan Association Rule Pada Faktor-Faktor Penentu Pemilihan Jurusan Menggunakan Random Forest." Penelitian ini didasarkan pada kebutuhan untuk memahami dinamika dan faktor-faktor yang mempengaruhi pemilihan jurusan di sekolah tersebut. SMK Negeri 4 Palembang menawarkan berbagai jurusan dengan karakteristik dan prospek yang berbeda-beda. Memahami faktor-faktor yang mempengaruhi keputusan calon siswa adalah langkah krusial untuk

membantu sekolah menyediakan informasi dan bimbingan yang efektif bagi calon siswa.

Metode Association Rule dipilih karena kemampuannya dalam mengidentifikasi hubungan mendalam antar faktor, yang dapat mengungkap pola-pola tersembunyi dalam data pemilihan jurusan. Teknik ini sangat berguna dalam situasi di mana berbagai faktor dapat saling berinteraksi dan mempengaruhi keputusan calon siswa secara kompleks. Association Rule memungkinkan kita melihat bagaimana faktor-faktor seperti prestasi akademik, minat, dan nilai saling terkait dalam mempengaruhi pemilihan jurusan.

Metode Random Forest dipilih karena keandalannya dalam memprediksi hasil berdasarkan data yang kompleks dan bervariasi. Algoritma ini mampu mengolah sejumlah besar data dan variabel, serta menghasilkan model prediksi yang akurat. Random Forest juga dapat menangani data yang tidak seimbang dan mengurangi risiko overfitting, sehingga hasil prediksi lebih dapat diandalkan. Kombinasi antara Association Rule dan Random Forest diharapkan memberikan pemahaman yang mendalam mengenai faktor-faktor yang mempengaruhi pemilihan jurusan, serta menghasilkan model prediksi yang akurat untuk mendukung pengambilan keputusan di SMK Negeri 4 Palembang.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, rumusan masalah dalam penelitian ini yaitu : “Bagaimana faktor-faktor penentu seperti nilai akademik (Matematika, IPA, Bahasa Inggris, dan Bahasa Indonesia) serta prestasi akademik dan non akademik dalam pemilihan jurusan calon siswa di SMK Negeri 4 Palembang berdasarkan metode Association Rule, dan model prediksi Random Forest dalam memberikan rekomendasi jurusan yang akurat serta relevan?”.

1.3 Tujuan dan Manfaat

1.3.1 Tujuan Penelitian

Adapun tujuan dari penelitian ini yaitu sebagai berikut :

1. Mengidentifikasi dan memahami faktor-faktor utama yang memengaruhi pemilihan jurusan calon siswa di SMK Negeri 4 Palembang menggunakan algoritma *Association Rule*, sehingga pihak sekolah dapat memperoleh wawasan tentang preferensi dan kecenderungan calon siswa.
2. Mengembangkan model prediksi pemilihan jurusan berbasis algoritma *Random Forest* yang dapat digunakan oleh sekolah untuk memberikan rekomendasi jurusan yang sesuai dengan minat dan potensi akademis calon siswa.

1.3.2 Manfaat Penelitian

Dengan adanya model prediksi pemilihan jurusan berbasis algoritma *Association Rule* dan *Random Forest*, diharapkan dapat memberikan informasi yang efektif digunakan oleh pihak sekolah untuk bimbingan karir calon siswa. Hal ini dapat membantu calon siswa dalam membuat keputusan yang lebih tepat, sesuai dengan minat dan potensi mereka, serta meningkatkan kepuasan dan keberhasilan akademik.

1.4 Batasan Masalah

Adapun batasan masalah dari penelitian ini ialah sebagai berikut :

1. Penelitian ini hanya dilakukan pada calon siswa SMK Negeri 4 Palembang dan tidak mencakup sekolah-sekolah lain di Palembang atau di daerah lainnya.
2. Data yang digunakan dalam penelitian ini terbatas pada data faktor-faktor penentu pemilihan jurusan yang tersedia di SMK Negeri 4 Palembang, seperti minat pribadi, prestasi akademik, dan nilai.
3. Penelitian ini menggunakan metode *Association Rule* dan *Random Forest*.
4. Data yang digunakan dalam penelitian ini adalah data historis yang dikumpulkan selama PPDB tahun ajaran 2024 / 2025.

1.5 Sistematika Penulisan

Sistematika penulisan ini disusun untuk mempermudah pemahaman terhadap pembahasan yang disajikan dalam laporan skripsi. Adapun sistematika penulisan ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab I Pendahuluan mencakup latar belakang mengenai pemilihan jurusan di SMK Negeri 4 Palembang, perumusan dan batasan masalah terkait faktor-faktor penentu pemilihan jurusan, serta data yang digunakan dalam penelitian. Selain itu, bab ini juga menjelaskan tujuan dan manfaat dari penerapan algoritma Association Rule dan Random Forest untuk analisis pemilihan jurusan, serta sistematika penulisan.

BAB II LANDASAN TEORI

Bab II Landasan Teori menguraikan teori-teori yang mendukung penelitian ini. Bab ini terdiri dari beberapa teori umum dan khusus yang relevan dengan pemecahan masalah pada penelitian, termasuk teori tentang Association Rule, Random Forest, dan faktor-faktor penentu pemilihan jurusan.

BAB III METODE PENELITIAN

Bab III Metode Penelitian membahas tahap-tahap penelitian yang meliputi studi literatur, pengumpulan data, pra-pemrosesan data, penerapan algoritma Association Rule dan Random Forest, serta pengujian metode menggunakan teknik validasi yang sesuai, seperti confusion matrix.

DAFTAR PUSTAKA

- Adlina Khairunnisa, Khairil Anwar Notodiputro, and Bagus Sartono. 2023. "A Study of Rule Extraction from Double Random Forest to Identify the Characteristics of Working Poor in Jakarta Province, Indonesia." *International Journal of Scientific Research in Science, Engineering and Technology* 258–66. doi: 10.32628/ijrsrset231069.
- Alduailij, Mona, Qazi Waqas Khan, Muhammad Tahir, Muhammad Sardaraz, Mai Alduailij, and Fazila Malik. 2022. "Machine-Learning-Based DDoS Attack Detection Using Mutual Information and Random Forest Feature Importance Method." *Symmetry* 14(6):1095. doi: 10.3390/sym14061095.
- Alfando, Alfando, and Regiolina Hayami. 2023. "KLASIFIKASI TEKS BERITA BERBAHASA INDONESIA MENGGUNAKAN MACHINE LEARNING DAN DEEP LEARNING: STUDI LITERATUR." *JATI (Jurnal Mahasiswa Teknik Informatika)* 7(1):681–86. doi: 10.36040/jati.v7i1.6486.
- Amini, Saeid, Mohsen Saber, Hamidreza Rabiei-Dastjerdi, and Saeid Homayouni. 2022. "Urban Land Use and Land Cover Change Analysis Using Random Forest Classification of Landsat Time Series." *Remote Sensing* 14(11):2654. doi: 10.3390/rs14112654.
- Bu'ulolo, Efori. 2024. "Algoritma K-Nearest Neighbor (K-NN) Dengan Normalisasi Max Min Untuk Menentukan Calon Mahasiswa Yang Layak Menerima KIP Kuliah Merdeka." *Simkom* 9(2):184–94. doi: 10.51717/simkom.v9i2.445.
- Crivei, Liana Maria, Gabriela Czibula, and Andrei Mihai. 2019. "A Study on Applying Relational Association Rule Mining Based Classification for Predicting the Academic Performance of Students." Pp. 287–300 in *Knowledge Science, Engineering and Management*, edited by C. Douligeris, D. Karagiannis, and D. Apostolou. Cham: Springer International Publishing.
- Dudek, Grzegorz. 2022. "A Comprehensive Study of Random Forest for Short-Term Load Forecasting." *Energies* 15(20):7547. doi: 10.3390/en15207547.

- Fernandez-Basso, Carlos, M. Dolores Ruiz, and Maria J. Martin-Bautista. 2024. "New Spark Solutions for Distributed Frequent Itemset and Association Rule Mining Algorithms." *Cluster Computing* 27(2):1217–34. doi: 10.1007/s10586-023-04014-w.
- Fister, Iztok, Iztok Fister, Dušan Fister, Vili Podgorelec, Iztok Fister, and Sancho Salcedo-Sanz. 2023. "A Comprehensive Review of Visualization Methods for Association Rule Mining: Taxonomy, Challenges, Open Problems and Future Ideas." *Expert Systems with Applications* 233. doi: 10.1016/j.eswa.2023.120901.
- Ilma, Hafizah, Khairil Anwar Notodiputro, and Bagus Sartono. 2023. "Association Rules in Random Forest for the Most Interpretable Model." *BAREKENG: Jurnal Ilmu Matematika Dan Terapan* 17(1):0185–96. doi: 10.30598/barekengvol17iss1pp0185-0196.
- Irie, Shinnosuke, Hirohisa Aman, Sousuke Amasaki, Tomoyuki Yokogawa, and Minoru Kawahara. 2023. "A Comparative Study of Hybrid Fault-Prone Module Prediction Models Using Association Rule and Random Forest." Pp. 33–38 in *2023 The 5th World Symposium on Software Engineering (WSSE)*. New York, NY, USA: ACM.
- Kim, Yesuel, and Youngchul Kim. 2022. "Explainable Heat-Related Mortality with Random Forest and SHapley Additive ExPlanations (SHAP) Models." *Sustainable Cities and Society* 79:103677. doi: 10.1016/j.scs.2022.103677.
- Kusnaldi, Muhammad Rafli, Timotius Gulo, and Soeb Aripin. 2022. "Penerapan Normalisasi Data Dalam Mengelompokkan Data Mahasiswa Dengan Menggunakan Metode K-Means Untuk Menentukan Prioritas Bantuan Uang Kuliah Tunggal." *Journal of Computer System and Informatics (JoSYC)* 3(4):330–38. doi: 10.47065/josyc.v3i4.2112.
- Lin, Wei Chun, Chien Hsiung Huang, Liang Tien Chien, Hsiao Jung Tseng, Chip Jin Ng, Kuang Hung Hsu, Chi Chun Lin, and Cheng Yu Chien. 2022. "Tree-Based Algorithms and Association Rule Mining for Predicting Patients' Neurological Outcomes After First-Aid Treatment for an Out-of-Hospital Cardiac Arrest During COVID-19 Pandemic: Application of Data Mining."

International Journal of General Medicine 15(September):7395–7405. doi: 10.2147/IJGM.S384959.

- Luchia, Nanda Try, Mustakim Mustakim, Noviarni Noviarni, Kelik Sussolaikah, and Teguh Arifianto. 2024. “Feature Selection In Support Vector Machine And Random Forest Algorithms For The Classification Of Recipients Of The Smart Indonesia Program.” Pp. 1–6 in *2024 International Conference on Circuit, Systems and Communication (ICCSC)*. IEEE.
- Matharaarachchi, Surani, Mike Domaratzki, Alan Katz, and Saman Muthukumarana. 2022. “Discovering Long COVID Symptom Patterns: Association Rule Mining and Sentiment Analysis in Social Media Tweets.” *JMIR Formative Research* 6(9):1–15. doi: 10.2196/37984.
- Moh’d, Fatma Hilali, Khairil Anwar Notodiputro, and Yenni Angraini. 2024. “Enhancing Interpretability in Random Forest: Leveraging InTrees for Association Rule Extraction Insights.” *IAES International Journal of Artificial Intelligence* 13(4):4054–61. doi: 10.11591/ijai.v13.i4.pp4054-4061.
- Permana, Inggih, and Febi Nur Salisah Salisah. 2022. “Pengaruh Normalisasi Data Terhadap Performa Hasil Klasifikasi Algoritma Backpropagation.” *Indonesian Journal of Informatic Research and Software Engineering (IJIRSE)* 2(1):67–72. doi: 10.57152/ijirse.v2i1.311.
- Radhika, K., S. Valarmathy, S. Selvarasu, K. Bashkaran, and C. Srinivasan. 2023. “Predictive Road Sign Maintenance Using Random Forest Regression and IoT Data.” Pp. 348–53 in *2023 International Conference on Sustainable Communication Networks and Application (ICSCNA)*. IEEE.
- Rahayu, Kharisma, Vindi Fitria, Dhini Septhya, Rahmaddeni Rahmaddeni, and Lusiana Efrizoni. 2023. “Klasifikasi Teks Untuk Mendeteksi Depresi Dan Kecemasan Pada Pengguna Twitter Berbasis Machine Learning.” *MALCOM: Indonesian Journal of Machine Learning and Computer Science* 3(2):108–14. doi: 10.57152/malcom.v3i2.780.
- Rai, Vishnu Kumar, Santonab Chakraborty, and Shankar Chakraborty. 2023. “Association Rule Mining for Prediction of Covid-19.” *Decision Making: Applications in Management and Engineering* 6(1):365–78. doi:

10.31181/dmame0317102022r.

- Rhodes, Jake S., Adele Cutler, and Kevin R. Moon. 2023. "Geometry- and Accuracy-Preserving Random Forest Proximities." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45(9):10947–59. doi: 10.1109/TPAMI.2023.3263774.
- Riskiyono, Fajar, and Deni Mahdiana. 2024. "Implementation of Random Forest Algorithm for Graduation Prediction." *Sinkron* 8(3):1662–70. doi: 10.33395/sinkron.v8i3.13750.
- Sarker, Iqbal H. 2021. "Machine Learning: Algorithms, Real-World Applications and Research Directions." *SN Computer Science* 2(3):1–21. doi: 10.1007/s42979-021-00592-x.
- Sayal, Anu, Janhvi Jha, Chaithra N, Veethika Gupta, Ashulekha Gupta, Omdeep Gupta, and Minakshi Memoria. 2023. "Neural Networks And Machine Learning." Pp. 58–63 in *2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA)*.
- Singh, Ajit. 2019. "Foundations of Machine Learning." *SSRN Electronic Journal* (112):1–8. doi: 10.2139/ssrn.3399990.
- Suryadewiansyah, Muhammad Kamil, Teja Endra, and Eng Tju. 2020. "Naïve Bayes Dan Confusion Matrix Untuk Efisiensi Analisa Intrusion Detection System Alert." *Jurnal Nasional Teknologi Dan Sistem Informasi* 8(2):81–88.
- Ulfa Windi Humaira, Nailia Nurul Hikmah, Rizki Anfanni Fahmi. 2019. "Identifikasi Demografi Dan Faktor Pemilihan." *At-Thullab Jurnal* 1:71–84.