

**SISTEM REKOMENDASI JURUSAN KULIAH SISWA SMA
BERDASARKAN NILAI RAPOR DENGAN ALGORITMA
NAIVE BAYES**

*Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika*



Oleh:

**Adrian Sutansaty
NIM : 09021382025164**

**Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2024**

LEMBAR PENGESAHAN SKRIPSI

SISTEM REKOMENDASI JURUSAN KULIAH SISWA SMA
BERDASARKAN NILAI RAPOR DENGAN
ALGORITMA NAIVE BAYES

Oleh :

Adrian Sutansaty

NIM : 09021382025164

Palembang, (14 Agustus 2024)

Pembimbing I



Kanda Jamuar Miraswan, M.T.
NIP 199001092019031012

Pembimbing II



Desty Rodiah, M.T.
NIP 198912212020122011

Mengetahui,
Ketua Jurusan Teknik Informatika,



Hadipurnawan Satria, M.Sc., Ph.D.
NIP 198004182020121001

TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI

Pada hari Jumat tanggal 2 Agustus 2024 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Adrian Sutansaty

NIM : 09021382025164

Judul : Sistem Rekomendasi Jurusan Kuliah Siswa SMA Berdasarkan Nilai Rapor Dengan Algoritma Naive Bayes

dan dinyatakan **LULUS**.

1. Ketua

Alvi Syahrini Utami, S.Si., M.Kom.

NIP 197812222006042003

2. Penguji 1

Al Farissi, S.Kom., M.Cs.

NIP 198512152014041001

3. Pembimbing 1

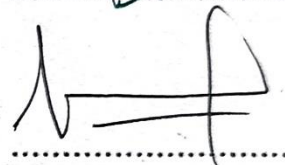
Kanda Januar Miraswan, M.T.

NIP 199001092019031012

4. Pembimbing 2

Desty Rodiah, M.T.

NIP 198912212020122011



Mengetahui,
Ketua Jurusan Teknik Informatika,

Hadipurnawan Satria, M.Sc., Ph.D.
NIP 198004182020121001

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Adrian Sutansaty
NIM : 09021382025164
Program Studi : Teknik Informatika
Judul Skripsi : SISTEM REKOMENDASI JURUSAN KULIAH SISWA SMA
BERDASARKAN NILAI RAPOR DENGAN ALGORITMA
NAIVE BAYES

Hasil Pengecekan Software *iThenticate/Turnitin* : 4%

Menyatakan bahwa laporan skripsi saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan skripsi 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, 30 Juli 2024



Adrian Sutansaty
NIM 09021382025164

Motto :

Algoritma Allah bekerja berbeda disini. Tidak bisa dijangkau dengan logika, dan terkadang tidak masuk akal. Jadi, kapanpun kemustahilan itu datang, berdoalah kepada Allah, dan berusaha. Kemustahilan adalah tempat bagi keajaiban Allah bekerja.

Kupersembahkan Karya Tulis Ini Kepada :

Allah S.W.T yang telah memberi kesehatan, kekuatan, dan kemudahan sehingga penulis mampu menyelesaikan skripsi ini

Ayah, Ibu dan Adik tercinta, yang senantiasa memberikan kasih sayang, dukungan, dan doa sehingga penulisan karya ilmiah ini bisa terselesaikan.

Dosen pembimbing skripsi, Bapak Kanda Januar Miraswan, M.T.
dan Ibu Desty Rodiah, M.T.

Kepala Sekolah SMA Negeri 1 Mesuji Raya, Bapak Huzairin, S.Pd., M.Si
Wakil Kepala Sekolah, Ibu Nova Heriyanti, S.Pd.,
Ibu Milatul Khasanah, S.Pd., Gr.

Almamater Fakultas Ilmu Komputer, Universitas Sriwijaya.

ABSTRACT

The determination of study programs for college is an essential for high school students who choose to pursue their studies. The aim of this research is to develop a web application-based college major recommendation system. This research applies a system design and development methodology using the Waterfall model approach. The data used is primary data which contains student report card scores taken over five semesters from the last several classes of SMAN 1 Mesuji Raya. System testing applies a combination of precision metric models and Leave-One-Out Cross Validation.

The results of this research produce a system that can recommend college majors and run as expected by applying the Naive Bayes and Euclidean Distance algorithms.

Keywords: Application, Naive Bayes, Euclidean Distance, System Development.

Palembang, (August 14, 2024)

Supervisor I



Kanda Januar Miraswan, M.T.
NIP 199001092019031012

Supervisor II



Desty Rodiah, M.T.
NIP 198912212020122011

Approved,
Head Department of Informatics Engineering,



Nadipurnawan Satria, M.Sc., Ph.D.
NIP 198004182020121001

ABSTRAK

Penentuan program studi di perguruan tinggi merupakan hal yang krusial bagi siswa Sekolah Menengah Atas (SMA) yang ingin melanjutkan studinya. Tujuan penelitian ini adalah mengembangkan sistem rekomendasi jurusan kuliah yang berbasis aplikasi web. Penelitian ini menerapkan metodologi desain dan pengembangan sistem dengan pendekatan model *Waterfall*. Data yang digunakan ialah data primer yang berisi nilai rapor siswa yang diambil selama lima semester dari beberapa angkatan terakhir dari SMAN 1 Mesuji Raya. Algoritma yang digunakan dalam sistem pengembangan ialah Naive Bayes dan Euclidean Distance. Pengujian sistem menerapkan kombinasi model metrik *precision* dan *Leave-One-Out Cross Validation*.

Hasil dari penelitian ini menghasilkan sistem yang dapat merekomendasikan jurusan kuliah dan berjalan sesuai harapan dengan menerapkan algoritma Naive Bayes dan Euclidean Distance.

Kata Kunci: Aplikasi, Naive Bayes, Euclidean Distance, Pengembangan Sistem.

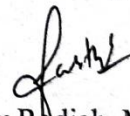
Pembimbing I



Kanda Januar Miraswan, M.T.
NIP 199001092019031012

Palembang, (14 Agustus 2024)

Pembimbing II



Desty Rediah, M.T.
NIP 198912212020122011

Mengetahui,
Ketua Jurusan Teknik Informatika,



Hadipurnawan Satria, M.Sc., Ph.D.
NIP 198004182020121001

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah melimpahkan rahmat, hidayah, serta karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Sistem Rekomendasi Jurusan Kuliah Siswa SMA Berdasarkan Nilai Rapor Dengan Algoritma Naive Bayes" sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Sriwijaya.

Dalam kesempatan ini, penulis ingin menyampaikan penghargaan dan ucapan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan, bimbingan, serta bantuan baik secara langsung maupun tidak langsung selama proses penulisan skripsi ini. Ucapan terima kasih khusus penulis sampaikan kepada:

1. Prof. DR. Erwin, S.Si., M.Si., Dekan Fakultas Ilmu Komputer Universitas Sriwijaya, yang telah memberikan fasilitas serta dukungan selama masa studi penulis.
2. Bapak Hadipurnawan Satria, M.Sc., Ph.D., Ketua Jurusan Teknik Informatika, yang telah memberikan izin dan dukungan dalam penyusunan skripsi ini.
3. Bapak Kanda Januar Miraswan, M.T., selaku pembimbing utama yang telah memberikan bimbingan, arahan, dan dukungan yang tiada henti dari awal hingga selesainya skripsi ini.
4. Ibu Desty Rodiah, M.T., selaku pembimbing kedua yang juga telah memberikan masukan berharga serta koreksi yang sangat membantu dalam penyempurnaan skripsi ini.
5. Kedua orang tua penulis, Bapak Maulana dan Ibu Rhiska Windyantie. yang dengan penuh kasih sayang dan kesabaran telah memberikan dukungan, doa, serta kasih sayang yang tiada henti dalam setiap langkah penulis hingga dapat menyelesaikan skripsi ini.

6. Kepala Sekolah SMA Negeri 1 Mesuji Raya, Bapak Huzairin, S.Pd., M.Si., Wakil Kepala Sekolah, Ibu Nova Heriyanti, S.Pd., Ibu Milatul Khasanah, S.Pd., Gr., beserta guru dan staff yang telah mengizinkan untuk menggunakan nilai rapor siswa sebagai objek utama dalam penelitian ini.
7. Kakak Agus Salim dan Ayuk Lienda Fatma Dewi. Terimakasih telah memberikan fasilitas dan bantuan yang berharga selama penulis menyusun skripsi ini.
8. Sahabat pejuang S.Kom, Retno Tri Aprilia, Hanny Putri Gayatri, Andelle Gianzra Basae, Febiyanti, dan Nadya Angelia, terimakasih atas motivasi, bantuan, dan dukungan yang telah banyak diberikan selama penulis melaksanakan perkuliahan hingga menyelesaikan skripsi.
9. Teman-teman serta rekan-rekan seperjuangan di Program Studi Teknik Informatika Universitas Sriwijaya, Silvi Agustina, Annisa Arrayyan, Aldo Haliansyah, Felixius Nilsen, Montelo, Farhan Revanza, M Zidane Kasfianto, Wahyu Safitra, serta teman-teman lainnya yang telah menjadi tempat berbagi ilmu dan pengalaman selama masa studi.
10. Semua pihak yang telah membantu penulis dalam penyelesaian skripsi ini, yang tidak dapat disebutkan satu per satu, namun kontribusinya sangat berarti bagi penulis.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis memohon maaf atas kesalahan dan kekurangan. Penulis juga sangat mengharapkan saran dan kritik yang membangun demi perbaikan di masa mendatang. Semoga skripsi ini dapat bermanfaat bagi para pembaca dan menjadi sumbangsih kecil bagi pengembangan ilmu pengetahuan.

Akhir kata, penulis berharap semoga Allah SWT selalu melimpahkan rahmat dan hidayah-Nya kepada kita semua.

Palembang, Agustus 2024

Penyusun,

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	ii
TANDA LULUS UJIAN KOMPREHENSIF	iii
HALAMAN PERNYATAAN.....	iv
MOTTO	v
ABSTRACT	vi
ABSTRAK	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
BAB I.....	I-1
PENDAHULUAN.....	I-1
1.1 Pendahuluan	I-1
1.2 Latar Belakang	I-1
1.3 Rumusan Masalah	I-3
1.4 Tujuan Penelitian	I-3
1.5 Manfaat Penelitian	I-4
1.6 Batasan Masalah.....	I-5
1.7 Sistematika Penulisan	I-6
1.8 Kesimpulan	I-7
BAB II	II-1
KAJIAN LITERATUR.....	II-1
2.1 Pendahuluan	II-1
2.2 Landasan Teori.....	II-1
2.2.1 Data Mining	II-1
2.2.2 Algoritma Naive Bayes	II-5
2.2.3 Euclidean Distance.....	II-6
2.2.4 Model Pengembangan Waterfall.....	II-7
2.2.5 Positive Predictive Value (PPV) dengan Precision.....	II-9
2.2.6 Leave-One-Out Cross Validation (LOOCV)	II-10
2.3 Penelitian Lain yang Relevan	II-12
2.3.1 Klasifikasi dan Rekomendasi Jurusan Kuliah Bagi Pelajar SMA Menggunakan Algoritme Naive Bayes-WP.....	II-12
2.3.2 Implementasi Data Mining dengan Algoritma Naive Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako	II-12
2.3.3 Implementasi Pengolahan Citra untuk Mendapatkan Informasi Mengenai Citra Baret TNI dengan Menggunakan Algoritma Euclidean Distance.....	II-13
2.4 Kesimpulan	II-15

BAB III.....	III-1
METODOLOGI PENELITIAN	III-1
3.1 Pendahuluan	III-1
3.2 Unit Penelitian.....	III-1
3.3 Teknik Pengumpulan Data	III-1
3.3.1 Jenis Data	III-1
3.3.2 Sumber Data.....	III-2
3.3.3 Metode Pengumpulan Data	III-2
3.4 Tahapan Penelitian	III-3
3.4.1 Kerangka Kerja	III-3
3.4.1.1 Tahap Pelatihan	III-4
3.4.1.2 Tahap Pengujian	III-5
3.4.2 Kriteria Pengujian	III-6
3.4.3 Format Data Pengujian.....	III-6
3.4.4 Alat yang Digunakan dalam Pelaksanaan Penelitian	III-6
3.4.5 Pengujian Penelitian.....	III-7
3.4.6 Analisis Hasil Pengujian	III-10
3.5 Metode Pengembangan Perangkat Lunak	III-12
3.6 Manajemen Proyek Penelitian.....	III-13
3.7 Kesimpulan.....	III-14
BAB IV.....	IV-1
PENGEMBANGAN PERANGKAT LUNAK.....	IV-1
4.1 Pendahuluan	IV-1
4.2 Analisis Kebutuhan	IV-1
4.3 Desain Sistem.....	IV-1
4.3.1 Arsitektur	IV-1
4.3.1.1 Use Case Diagram	IV-2
4.3.1.2 Activity Diagram.....	IV-3
4.3.1.3 Sequence Diagram.....	IV-8
4.3.1.4 Class Diagram	IV-15
4.4 Implementasi	IV-18
4.4.1 Spesifikasi Perangkat	IV-18
4.4.2 Bahasa Pemrograman.....	IV-19
4.4.3 Implementasi Antarmuka.....	IV-19
4.5 Pengujian Sistem	IV-26
4.5.1 Pengujian Black Box.....	IV-26
4.5.2 Rencana Pengujian.....	IV-26
4.5.3 Hasil Pengujian	IV-29
4.6 Kesimpulan.....	IV-32
BAB V	V-1
HASIL DAN PEMBAHASAN	V-1

5.1 Pendahuluan	V-1
5.2 Deskripsi Dataset.....	V-1
5.3 Proses Pengujian.....	V-2
5.3.1 Metode Pengujian Pertama (Menggunakan Keseluruhan Dataset).....	V-4
5.3.2 Metode Pengujian Kedua (Menggunakan Dataset dengan Minimal Terdapat 2 Data per Jurusan)	V-7
5.4 Analisis Hasil Pengujian.....	V-9
5.5 Kesimpulan.....	V-10
BAB VI.....	VI-1
KESIMPULAN DAN SARAN	VI-1
6.1 Pendahuluan	VI-1
6.2 Kesimpulan.....	VI-1
6.3 Saran.....	VI-2
DAFTAR PUSTAKA.....	xvi
LAMPIRAN.....	xix

DAFTAR TABEL

Tabel III-1. Standar Konversi Nilai	III-8
Tabel III-2. Model <i>Leave-One-Out Cross Validation</i>	III-10
Tabel IV-1. Spesifikasi perangkat keras	IV-18
Tabel IV-2. Spesifikasi perangkat lunak.....	IV-18
Tabel IV-3. Rencana Pengujian	IV-26
Tabel IV-4. Hasil Pengujian.....	IV-29
Tabel V-1. Hasil Pengujian Pertama.....	V-4
Tabel V-2. Hasil Pengujian Kedua.....	V-7

DAFTAR GAMBAR

Gambar II-1. Proses <i>Data Mining</i> (Han dkk., 2006).....	II-3
Gambar II-2. Skema <i>Leave-One-Out Cross Validation</i>	II-10
Gambar III-1. Kerangka Kerja	III-3
Gambar III-2. Contoh Penghitungan Nilai Rata-Rata.....	III-4
Gambar III-3. Contoh Proses Membuat Model Naive Bayes	III-5
Gambar III-4. Contoh <i>Additive Smoothing</i>	III-9
Gambar IV-1. <i>Use Case Diagram</i>	IV-2
Gambar IV-2. Menggunakan Sistem Rekomendasi Jurusan Kuliah.....	IV-3
Gambar IV-3. <i>Admin</i> Melakukan Login	IV-4
Gambar IV-4. <i>Admin</i> Melihat Data Siswa	IV-4
Gambar IV-5. <i>Admin</i> Melihat Data Jurusan dan Kampus.....	IV-5
Gambar IV-6. <i>Admin</i> Melihat Data Siswa <i>Eligible</i>	IV-5
Gambar IV-7. <i>Admin</i> Melakukan <i>Update Dataset</i>	IV-6
Gambar IV-8. <i>Admin</i> Melakukan Pengujian.....	IV-6
Gambar IV-9. <i>Admin</i> Melihat Riwayat dan <i>Log</i> Pengujian	IV-7
Gambar IV-10. <i>Admin</i> Menghapus Riwayat Pengujian.....	IV-7
Gambar IV-11. Sequence Diagram : Menjalankan Sistem Rekomendasi	IV-8
Gambar IV-12. <i>Sequence Diagram</i> : Melakukan <i>Login</i> ke Dasbor Admin.....	IV-9
Gambar IV-13. <i>Sequence Diagram</i> : Melihat Data Siswa	IV-10
Gambar IV-14. <i>Sequence Diagram</i> : Melihat Data Jurusan & Universitas	IV-10
Gambar IV-15. <i>Sequence Diagram</i> : Melihat Data Siswa <i>Eligible</i>	IV-10
Gambar IV-16. <i>Sequence Diagram</i> : Mengunggah Data Siswa	IV-11
Gambar IV-17. <i>Sequence Diagram</i> : Membuat Data Latih	IV-12
Gambar IV-18. <i>Sequence Diagram</i> : Melakukan Pengujian.....	IV-13
Gambar IV-19. <i>Sequence Diagram</i> : Melihat Riwayat dan <i>Log</i> Pengujian.....	IV-14
Gambar IV-20. <i>Sequence Diagram</i> : Menghapus Riwayat Pengujian	IV-14
Gambar IV-21. <i>Class Diagram</i>	IV-15
Gambar IV-22. Halaman Login	IV-19
Gambar IV-23. Halaman Dasbor	IV-20
Gambar IV-24. Halaman Lihat Data : Semua Siswa	IV-20
Gambar IV-25. Halaman Lihat Data : Jurusan & Kampus	IV-21
Gambar IV-26. Halaman Lihat Data : Siswa <i>Eligible</i>	IV-21
Gambar IV-27. Halaman <i>Update Dataset</i>	IV-22
Gambar IV-28. Halaman Rekomendasi Jurusan.....	IV-22
Gambar IV-29. Hasil Rekomendasi Jurusan.....	IV-23
Gambar IV-30. Keterangan Warna Tabel Pada Hasil.....	IV-23
Gambar IV-31. Halaman Pengujian.....	IV-24
Gambar IV-32. Menampilkan Hasil Pengujian.....	IV-24

Gambar IV-33. Halaman Riwayat Pengujian.....	IV-25
Gambar IV-34. Halaman Log Riwayat Pengujian	IV-25
Gambar V-1. <i>Template Dataset</i>	V-1
Gambar V-2. Hasil Pengujian Pertama	V-4
Gambar V-3. Hasil Pengujian Kedua.....	V-7

BAB I

PENDAHULUAN

1.1 Pendahuluan

Pada bab ini akan digambarkan latar belakang permasalahan, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah, serta sistematika dalam penelitian Sistem Rekomendasi Jurusan Kuliah Siswa SMA Berdasarkan Nilai Rapor Dengan Algoritma Naive Bayes.

1.2 Latar Belakang

Seleksi nasional masuk perguruan tinggi negeri (PTN) pada SNBP dan SPAN-PTKIN merupakan jalur masuk PTN yang paling diharapkan bagi siswa SMA yang berencana melanjutkan pendidikan ke perguruan tinggi. Pasalnya, jalur ini dapat menjadi penentu dari keberlanjutan akademik seorang siswa, terutama bagi mereka yang memiliki kendala finansial untuk masuk perguruan tinggi. Mereka yang berhasil lolos melalui jalur ini dapat lebih menghemat biaya pendidikan karena mendapatkan beasiswa KIP dari pemerintah (Nizam, 2020). Namun sayangnya, tidak sedikit siswa SMA yang dicap sebagai siswa unggul dan memiliki nilai rapor yang cukup bagus, namun tidak lolos masuk jurusan yang dipilihnya (Bintang, 2020).

Ada beberapa faktor yang mempengaruhi lolos atau tidaknya seseorang pada seleksi nasional masuk PTN, yaitu nilai rapor siswa, prestasi siswa, serta akreditasi sekolah (Alfian, 2023). Ketidaksesuaian dalam pemilihan jurusan juga menjadi faktor yang menyebabkan tidak lolosnya siswa dalam seleksi nasional masuk PTN, seperti tuntutan orang tua, gengsi, kurang pemahaman terhadap kecocokan jurusan dengan minat bakat diri, ataupun sekedar ikut-ikutan (Fahmi, 1997).

SMAN 1 Mesuji Raya merupakan salah satu dari beberapa SMA di Indonesia yang memiliki permasalahan serupa. Tidak sedikit dari siswa mereka yang mendapatkan kesempatan melalui seleksi masuk PTN tanpa tes, namun pada akhirnya mendapatkan hasil yang tidak sesuai harapan. Umumnya, permasalahan ini terjadi karena ketidaksesuaian jurusan dengan minat bakat diri yang dipresentasikan dari nilai rapor siswa. Akibatnya, tingkat kelulusan siswa SMAN 1 Mesuji Raya pada seleksi masuk PTN tanpa tes cenderung rendah. Dari data yang didapatkan, pada angkatan terakhir (tahun 2023) terdapat total 97 siswa dengan hanya 8 orang yang berhasil lolos seleksi nasional masuk PTN melalui jalur SNBP ataupun SPAN-PTKIN.

Pemilihan jurusan kuliah ini sudah seharusnya dipertimbangkan sedini mungkin oleh pelajar SMA. Sangat penting bagi calon mahasiswa untuk mengetahui minat dan bakat yang dimilikinya agar dapat menentukan jurusan kuliah yang cocok.

Berdasarkan latar belakang diatas, maka peneliti mengambil judul “**Sistem Rekomendasi Jurusan Kuliah Siswa SMA Berdasarkan Nilai Rapor Dengan Algoritma Naive Bayes**”. Penelitian ini memiliki urgensi yang tinggi karena hasil klasifikasi dapat dimanfaatkan untuk meningkatkan potensi kelulusan siswa pada seleksi masuk PTN tanpa tes. Sistem rekomendasi dilakukan dengan mengklasifikasi nilai rapor siswa peserta jalur masuk PTN tanpa tes dengan siswa yang sudah berhasil lolos beberapa periode sebelumnya.

1.3 Rumusan Masalah

Berdasarkan latar belakang yang sudah dikumpulkan, dapat dirumuskan beberapa permasalahan sebagai berikut:

1. Bagaimana cara mengembangkan sistem untuk menentukan jurusan kuliah yang sesuai berdasarkan nilai rapor siswa SMA?
2. Bagaimana sistem rekomendasi dapat memberikan informasi tentang mata pelajaran yang masih perlu dimaksimalkan dalam kegiatan belajar mengajar di SMA terkait?
3. Bagaimana cara mendapatkan efektivitas rekomendasi jurusan kuliah yang sesuai dengan jumlah dataset yang minim?
4. Bagaimana mengatasi potensi bias pada hasil rekomendasi jurusan?

Penelitian ini bertujuan untuk menjawab permasalahan tersebut dan memberikan kontribusi dalam meningkatkan potensi kelulusan dalam seleksi perguruan tinggi negeri.

1.4 Tujuan Penelitian

Tujuan utama penelitian ini adalah mengembangkan sistem rekomendasi kuliah berdasarkan nilai rapor siswa di SMAN 1 Mesuji raya melalui implementasi algoritma Naive Bayes berbasis aplikasi web. Secara spesifik, tujuan dari penelitian ini adalah:

1. Mengembangkan model klasifikasi dengan data sampel yang sangat minim melalui pertimbangan variabel-variabel kinerja akademik siswa, seperti nilai matematika, bahasa Inggris, sains, dan lainnya, untuk merumuskan rekomendasi jurusan kuliah.
2. Memberikan informasi tentang mata pelajaran mana saja yang perlu dimaksimalkan oleh siswa.
3. Menguji keakuratan dan efektivitas model klasifikasi yang telah dikembangkan menggunakan data siswa angkatan sebelumnya.
4. Mengoptimalkan model rekomendasi jurusan kuliah berdasarkan hasil evaluasi.

1.5 Manfaat Penelitian

1. Bagi Siswa SMA

Siswa SMA dapat memperoleh rekomendasi jurusan kuliah yang sesuai dengan kemampuan dan minat akademik mereka, membantu mereka membuat keputusan yang lebih cerdas dalam pemilihan jurusan kuliah. Dengan memilih jurusan yang sesuai, siswa memiliki peluang kelulusan yang lebih tinggi pada jalur seleksi nasional masuk PTN.

2. Bagi Institusi Pendidikan

Institusi pendidikan dapat meningkatkan kualitas kelulusan siswa dengan memberikan bimbingan karir yang lebih efektif dan akurat, membantu siswa memilih jurusan yang sesuai dengan potensi mereka. Hasilnya, institusi pendidikan dapat meningkatkan reputasi mereka dengan menghasilkan lulusan berkualitas, sesuai dengan kebutuhan dan tuntutan pasar kerja.

3. Bagi Orang Tua Siswa

Orang tua siswa mendapatkan pemahaman yang lebih mendalam tentang minat dan potensi akademik anak-anak mereka, membantu mereka memberikan dukungan yang lebih baik dalam pemilihan jurusan kuliah.

4. Bagi Perguruan Tinggi Negeri

Perguruan tinggi negeri mendapatkan mahasiswa yang lebih berkualitas dan sesuai dengan program studi yang mereka pilih, mengurangi risiko *drop-out* dan peningkatan prestasi akademik.

5. Bagi Dunia Pendidikan dan Penelitian

Penelitian ini dapat memberikan referensi dalam pengembangan penelitian *machine learning* yang menggunakan algoritma Naive Bayes sebagai algoritma utamanya.

6. Bagi Masyarakat dan Ekonomi

Dengan memiliki lulusan yang lebih sesuai dengan kebutuhan pasar kerja, maka terjadilah peningkatan produktivitas dan kompetensi di berbagai sektor industri yang secara langsung dapat mengurangi angka pengangguran karena peningkatan kesesuaian antara lulusan dengan sumber daya manusia yang ada.

Dengan memberikan manfaat ini, penelitian ini diharapkan tidak hanya memberikan kontribusi dalam bidang pendidikan dan pemilihan karier, tetapi juga dalam memperkuat kualitas lulusan yang dapat beradaptasi dengan cepat dan efisien dengan tuntutan pasar kerja yang terus berkembang.

1.6 Batasan Masalah

Merujuk kepada informasi sebelumnya, untuk mencegah terjadinya bias atau perluasan isu yang tidak diinginkan, peneliti memutuskan untuk membatasi cakupan masalah dalam penelitian ini sebagai berikut:

1. Penelitian ini terbatas pada siswa SMAN 1 Mesuji Raya, Indonesia. Namun, sistem rekomendasi tetap dapat digunakan secara maksimal oleh SMA lainnya.
2. Populasi penelitian melibatkan siswa SMAN 1 Mesuji Raya yang ingin meneruskan ke jenjang kuliah melalui jalur seleksi nasional masuk PTN.
3. Data yang digunakan terbatas pada siswa yang berhasil lolos melalui jalur masuk tanpa tes pada periode sebelumnya. Data ini mencakup nilai rapor, hasil seleksi masuk PTN tanpa tes.
4. Nilai rapor siswa SMA di masa mendatang dapat digunakan untuk menambahkan jumlah *dataset* dengan tujuan meningkatkan jumlah variasi jurusan dan akurasi dari algoritma. Namun, penelitian ini dibatasi hingga sistem dapat memberi keluaran yang berupa rekomendasi jurusan, nilai mata pelajaran yang masih perlu dimaksimalkan, beserta rekomendasi jurusan yang ada pada data latih saja.

1.7 Sistematika Penulisan

BAB I. PENDAHULUAN

Bab ini berisi uraian mengenai latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah, metodologi penelitian, dan sistematika penulisan.

BAB II. KAJIAN LITERATUR

Bab ini membahas dasar-dasar teori yang diterapkan dalam penelitian, seperti teori *Data Mining*, algoritma Naive Bayes, Euclidean Distance, model pengembangan *Waterfall*, *Positive Predictive Value* dengan *Precision*, *Leave-One-Out Cross Validation*, penelitian lain yang relevan, serta kesimpulan dari pembahasan tersebut.

BAB III. METODOLOGI PENELITIAN

Bab ini menjabarkan tahapan-tahapan yang dilakukan dalam penelitian. Tahapan penelitian akan dideskripsikan secara rinci dengan perencanaan yang mengacu pada kerangka kerja, dan *rundown* penelitian yang mengacu pada manajemen proyek penelitian pada akhir bab ini.

1.8 Kesimpulan

Penelitian ini bermula dari pemahaman bahwa seleksi nasional masuk perguruan tinggi negeri (PTN) tanpa tes, seperti SNBP dan SPAN-PTKIN, merupakan peluang besar bagi siswa SMA, terutama yang mengalami kendala finansial. Meskipun demikian, banyak siswa SMA yang mengalami kesulitan dalam memilih jurusan kuliah yang sesuai dengan minat dan bakatnya, sehingga berdampak pada tingkat kelulusan yang rendah.

Penelitian ini akan mengimplementasikan Naive Bayes sebagai algoritma utama untuk membantu siswa SMA, dalam menentukan jurusan kuliah yang sesuai berdasarkan nilai rapornya.

Rumusan masalah yang diajukan mencakup pertanyaan tentang efektivitas siswa dalam memilih jurusan, faktor-faktor yang mempengaruhi ketidaksesuaian jurusan dengan minat bakat siswa, dan bagaimana sistem rekomendasi jurusan kuliah dapat membantu dalam memberikan rekomendasi jurusan yang sesuai. Tujuan penelitian ini mencakup pengembangan sistem rekomendasi jurusan kuliah untuk siswa SMA, dengan harapan dapat meningkatkan kesesuaian jurusan dengan minat dan bakat siswa.

Kemudian, manfaat penelitian disajikan untuk berbagai pihak yang mencakup peningkatan kesesuaian jurusan dengan kebutuhan pasar kerja, pengurangan risiko drop-out, peningkatan kualitas lulusan, dan kontribusi pada pengembangan *machine learning* dengan algoritma Naive Bayes.

Batasan-batasan penelitian juga disajikan untuk memastikan fokus dan kejelasan penelitian. Batasan-batasan tersebut mencakup aspek geografis, populasi penelitian, variabel yang ditentukan, sampel penelitian, dan keterbatasan data.

Dengan demikian, bab pendahuluan ini memberikan landasan yang kuat untuk melanjutkan penelitian dengan fokus untuk mengembangkan sistem rekomendasi jurusan kuliah untuk siswa SMA dengan Naive Bayes sebagai algoritma utamanya.

DAFTAR PUSTAKA

- Damuri, A., Riyanto, U., Rusdianto, H., & Aminudin, M. (2021). Implementasi Data Mining dengan Algoritma Naive Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako. *JURIKOM (Jurnal Riset Komputer)/Jurikom*, 8(6), 219. <https://doi.org/10.30865/jurikom.v8i6.3655>
- Douglass, M. J. J. (2020). Book Review: Hands-on Machine Learning with Scikit-Learn, Keras, and Tensorflow, 2nd edition by Aurélien Géron. *Physical and Engineering Sciences in Medicine/Physical and Engineering Sciences in Medicine*, 43(3), 1135–1136. <https://doi.org/10.1007/s13246-020-00913-z>
- Fitriawanti, R., Cholissodin, I., & Dewi, R. K. (2018). Klasifikasi dan Rekomendasi Jurusan Kuliah Bagi Pelajar SMA Menggunakan Algoritme Naive Bayes-WP. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(11), 4914–4922. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/3129>
- John, G. H., & Langley, P. (2013, February 20). *Estimating Continuous Distributions in Bayesian Classifiers*. arXiv.org. <https://arxiv.org/abs/1302.4964>

- Khadka, N., & Khadka, N. (2023, December 6). How Leave-One-Out Cross Validation (LOOCV) Improve's model performance - Dataaspirant. *Dataaspirant - A Data Science Portal For Beginners*.
<https://dataaspirant.com/leave-one-out-cross-validation-loocv/>
- Mitchell, M. (1998). *An introduction to genetic algorithms*.
<https://doi.org/10.7551/mitpress/3927.001.0001>
- Powers, D. M. W. (2020, October 11). *Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation*. arXiv.org.
<https://arxiv.org/abs/2010.16061>
- Pratidina, A., & Affandi, H. (2017). IMPLEMENTASI PENGOLAHAN CITRA UNTUK MENDAPATKAN INFORMASI MENGENAI CITRA BARET TNI DENGAN MENGGUNAKAN ALGORITMA EUCLIDEAN DISTANCE. *Pratidina | Jurnal Ilmiah Teknologi Dan Rekayasa*.
<https://ejournal.gunadarma.ac.id/index.php/tekno/article/view/1602>
- Ridwan, M., Suyono, H., & Sarosa, M. (2013). Penerapan data mining untuk evaluasi kinerja akademik mahasiswa menggunakan algoritma Naive Bayes Classifier. *Jurnal EECCIS*, 7(1), 59–64.
<https://jurnaleeccis.ub.ac.id/index.php/eccis/article/download/204/176>
- Santosa, B. (2011). *DATA MINING : Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta : Garah Ilmu.

- Saputro, I. W., & Sari, B. W. (2020). Uji Performa Algoritma Naive Bayes untuk Prediksi Masa Studi Mahasiswa. *Citec (Creative Information Technology) Journal/Citec Journal*, 6(1), 1. <https://doi.org/10.24076/citec.2019v6i1.178>
- Sethi, A. (2020, April 14). *Supervised Learning vs. Unsupervised Learning – A Quick Guide for Beginners*. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2020/04/supervised-learning-unsupervised-learning/>
- Sommerville, I. (2011). *Software Engineering, 9/e*. Pearson Education India.
- Vashisht, R. (2024, April 15). *How to solve the zero-frequency problem in Naive Bayes?* Atoti Community. <https://www.atoti.io/articles/how-to-solve-the-zero-frequency-problem-in-naive-bayes/>
- Zhang, H. (2004). The Optimality of Naive Bayes. *Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference*, 562–567.
- Zubaidi, M. M. (2016). *PENYESUAIAN DIRI PADA MAHASISWA SALAH JURUSAN*. <https://digilib.uin-suka.ac.id/id/eprint/23237>