LAPORAN KERJA PRAKTIK

Analisis Kecenderungan Siswa SMP/MTs Sederajat di Provinsi DI Yogyakarta dalam Pemilihan Tingkatan Akhir Pendidikan Menengah Menggunakan Metode Klasifikasi Decision Tree Algoritma C5.0



Disusun Oleh: Silvy Aqila Maharani 21/478973/PA/20777

Dosen Pembimbing:
Dr. Abdurakhman, S.Si., M.Si.
NIP. 197207041998031001

PROGRAM STUDI STATISTIKA
DEPARTEMEN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS GADJAH MADA
2023

LEMBAR PENGESAHAN LAPORAN KERJA PRAKTIK

Analisis Kecenderungan Siswa SMP/MTs Sederajat di Provinsi DI Yogyakarta dalam Pemilihan Tingkatan Akhir Pendidikan Menengah Menggunakan Metode Klasifikasi Decision Tree Algoritma C5.0

> Disusun Oleh: Silvy Aqila Maharani 21/478973/PA/20777

Telah Diperiksa dan Disetujui oleh Dosen Pembimbing

Yogyakarta, 2023

Menyetujui,

Dosen Pembimbing

Dr. Abdurakhman, S.Si., M.Si. NIP. 197207041998031001

PERNYATAAN

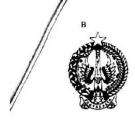
Dengan ini saya menyatakan bahwa laporan kerja praktik dengan judul "Analisis Kecenderungan Siswa SMP/MTs Sederajat di Provinsi DI Yogyakarta dalam Pemilihan Tingkatan Akhir Pendidikan Menengah Menggunakan Metode Klasifikasi Decision Tree Algoritma C5.0" tidak pernah diajukan sebelumnya untuk memperoleh gelar sarjana di suatu Perguruan Tinggi manapun dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 18 Desember 2023

Silvy Aqila Maharani

NIM 21/478973/PA/20777

SURAT KETERANGAN KERJA PRAKTIK



PEMERINTAH DAERAH DAERAH ISTIMEWA YOGYAKARTA DINAS PENDIDIKAN, PEMUDA, DAN OLAHRAGA

Jalan Cendana 9 Yogyakarta Telepon (0274) 550330, 513348 Faksimile (0274) 513348 Laman: <u>www.dikpora.jogjaprov.go.id</u> Email: <u>dikpora@jogjaprov.go.id</u> Kode Pos 55166

Yogyakarta, 08 Juni 2023

Nomor Sifat

070/05803

BIASA

Lamp Perihal

Ijin Kerja Praktik

Kepada Dekan Fakultas Matematika dan IPA Universitas Gajah Mada Yogyakarta

S.E., M.Pd. 4

198602 1 002

Yogyakarta

Menindaklanjuti surat dari Dekan Fakultas Matematika dan IPA Universitas Gajah Mada Yogyakarta Nomor : 200/UNI/FMIPA.1.1/AKD/PK.01.06/2023 tanggal 24 Mei 2023 perihal sesuai pada pokok surat, dengan ini kami sampaikan bahwa pada prinsipnya kami tidak berkeberatan dan mengijinkan mahasiswa :

Nama : Silvy Aqila Maharani

NIM

: 21/478973/PA/20777

Prodi

: Statistika

untuk melaksanakan program kerja praktik mulai tanggal 26 Juni 2023 s.d. 28 Juli 2023 di lingkungan Dinas Pendidikan, Pemuda dan Olahraga DIY.

DISDIKPOR

DIDIK WA

Atas perhatiannya kami ucapkan terima kasih.

KATA PENGANTAR

Dengan nama Allah Yang Maha Pengasih dan Maha Penyayang, puji syukur saya panjatkan ke hadirat-Nya atas segala rahmat, hidayah, serta karunia-Nya yang senantiasa melimpah. Shalawat serta salam semoga senantiasa terlimpahkan kepada Rasulullah Muhammad SAW, keluarga, sahabat, dan seluruh umat Islam yang setia.

Laporan analisis ini disusun sebagai hasil upaya saya dalam mengkaji dan memahami permasalahan yang tersebut dalam judul. Laporan ini mencerminkan dedikasi dan kerja keras saya dalam menganalisis, mengevaluasi, dan menyajikan informasi yang relevan, dengan tujuan untuk memberikan pemahaman yang lebih mendalam terhadap subjek yang menjadi fokus kajian.

Analisis ini tidak terlepas dari dukungan dan bimbingan berbagai pihak yang dengan ikhlas memberikan kontribusi pemikiran dan masukan berharga. Saya mengucapkan terima kasih yang sebesar-besarnya kepada rekan-rekan saya pada saat melaksanakan kerja praktik di Dinas Pendidikan, Pemuda, dan Olahraga Provinsi DI Yogyakarta yang senantiasa memberikan arahan, kritik, dan saran yang membangun. Tidak lupa, rasa terima kasih saya sampaikan kepada teman-teman yang turut serta dalam diskusi, memberikan inspirasi, serta semangat bagi kemajuan penelitian ini. Keterlibatan mereka menjadi pendorong semangat bagi saya untuk memberikan yang terbaik dalam penyusunan laporan ini. Serta seluruh pihak yang turut serta, baik secara langsung maupun tidak langsung, dalam perjalanan penulisan laporan ini saya sampaikan penghargaan yang setinggi-tingginya.

Akhir kata, doa saya semoga laporan ini dapat menjadi bahan referensi yang bermanfaat, membuka wawasan, dan menginspirasi bagi semua pihak yang berkepentingan. Saya menyadari sepenuhnya bahwa setiap karya merupakan perjalanan yang panjang, dan semoga karya ini dapat terus menjadi pijakan bagi perbaikan dan peningkatan di masa mendatang.

Yogyakarta, 18 Desember 2023

Silvy Aqila Maharani

NIM 21/478973/PA/20777

DAFTAR ISI

LAPORAN KERJA PRAKTIK	i
LEMBAR PENGESAHAN	ii
PERNYATAAN	iii
SURAT KETERANGAN KERJA PRAKTIK	iv
KATA PENGANTAR	v
DAFTAR ISI	vi
DAFTAR TABEL	viii
DAFTAR GAMBAR	ix
ABSTRAK	X
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	1
1.3 Tujuan Kerja Praktik	2
1.3.1 Tujuan Umum	2
1.3.2 Tujuan Khusus	2
1.4 Manfaat Kerja Praktik	2
1.4.1 Bagi Mahasiswa	2
1.4.2 Bagi Program Studi Sarjana Statistika	2
1.4.3 Bagi Instansi/Perusahaan	3
1.5 Waktu dan Tempat Pelaksanaan	3
1.5.1 Waktu Pelaksanaan	3
1.5.2 Tempat Pelaksanaan	3
1.6 Metodologi Penelitian	3
1.6.1 Metode Pengumpulan Data	3
1.6.2 Metode Analisis Data	3
BAB II PROFIL DINAS PENDIDIKAN, PEMUDA, DAN OLAHRAGA DIY	4
2.1 Visi dan Misi	4
2.1.1 Visi	4
2.1.2 Misi	4
2.2 Tujuan dan Sasaran	4
2.2.1 Tujuan	4
2.2.2 Sasaran	4

2.3 Struktur Organisasi	4
2.4 Tugas dan Fungsi	5
2.4.1 Tugas	I
2.4.2 Fungsi	, I
BAB III LANDASAN TEORI	6
3.1 Penerimaan Peserta Didik Baru (PPDB)	6
3.1.1 PPDB Jenjang Sekolah Menengah di Provinsi DIY6)
3.2 Decision Tree	7
3.2.1 Algoritma C5.0	!
3.3 Confusion Matrix	8
BAB IV HASIL DAN PEMBAHASAN	10
4.1 Deskripsi Data	10
4.2 Pre-Processing	10
4.3 Hasil Analisis	11
4.3.1 Decision Tree Algoritma C5.0	,
4.3.2 Confusion Matrix	i
BAB V PENUTUP	15
5.1 Kesimpulan	15
5.2 Saran	15
DAFTAR PUSTAKA	17
I AMDIDAN	10

DAFTAR TABEL

Tabel 1 Tabel Confusion Matrix	8
Tabel 2 Tabel Transformasi Data	10
Tabel 3 Table Nilai Entropy dan Gain	12
Tabel 4 Tabel Hasil Confusion Matrix dari Decision Tree Algoritma C5.0	13

DAFTAR GAMBAR

Bagan 1 Ba	gan Struktur Orga	nisasi Dinas Pe	ndidikan,	Pemuda,	dan Olahraga	DIY	. 5
Bagan 2 Ba	gan Hasil Decisio	n Tree Algoritn	na C5.0				13

ABSTRAK

Pendidikan menengah merupakan fase krusial dalam perjalanan pendidikan seseorang, di mana siswa dihadapkan pada pilihan antara Sekolah Menengah Atas (SMA) dan Sekolah Menengah Kejuruan (SMK). Penelitian ini dilakukan dengan menggunakan data PPDB SMA dan SMK di Provinsi DI Yogyakarta. Variabel-variabel yang dianalisis melibatkan aspek-aspek seperti tempat tinggal siswa, jenis kelamin, serta nilai ASPD. Metode klasifikasi Decision Tree Algoritma C5.0 digunakan untuk mengidentifikasi pola dan hubungan antara variabel-variabel tersebut pada pilihan siswa. Penelitian ini dapat memberikan kontribusi untuk pemahaman lebih mendalam terkait faktor-faktor yang memengaruhi keputusan siswa dalam pemilihan sekolah menengah, serta dapat menjadi landasan bagi pengambilan keputusan pendidikan di tingkat kebijakan. Dengan demikian, diharapkan hasil penelitian ini dapat membantu meningkatkan kualitas pendidikan di Provinsi DI Yogyakarta dan memberikan arahan bagi siswa dalam memilih jalur pendidikan yang sesuai dengan potensi dan tujuan mereka.

Kata Kunci: Decision Tree, Algoritma C5.0, Klasifikasi, SMA, SMK, PPDB

BABI

PENDAHULUAN

1.1 Latar Belakang

Dalam beberapa tahun terakhir, observasi terhadap pilihan sekolah lanjutan di kalangan siswa SMP menunjukkan tren di mana lebih banyak siswa memilih mendaftar di SMA dibandingkan Sekolah Menengah Kejuruan (SMK). Beberapa faktor dapat menjelaskan tren ini, termasuk pandangan masyarakat terhadap SMA dan SMK.

Siswa-siswa pada umumnya lebih tertarik untuk melanjutkan pendidikan mereka di SMA, mungkin karena persepsi bahwa SMA memberikan fondasi pendidikan yang lebih luas dan umum. SMA sering dianggap sebagai lembaga pendidikan yang menyediakan landasan akademis yang kuat, mempersiapkan siswa untuk mengikuti ujian nasional, dan memungkinkan akses ke perguruan tinggi.

Di sisi lain, pandangan terhadap SMK cenderung dipengaruhi oleh stereotip yang tidak selalu akurat. Sebuah stereotip yang masih melekat dalam masyarakat adalah bahwa SMK dihuni oleh siswa "kurang pintar" atau mereka yang tidak mampu mencapai prestasi akademis yang tinggi. Namun, pandangan ini tidak mempertimbangkan nilai penting dari SMK dalam mengembangkan keterampilan praktis dan keahlian industri yang sangat dihargai di dunia kerja.

Anak-anak yang sering dianggap "pintar" dalam konteks akademis mungkin merasa tekanan untuk memilih SMA sebagai pilihan lanjutan mereka. Hal ini dapat disebabkan oleh pandangan sosial yang masih mengutamakan prestise akademis dan keberhasilan ujian standar sebagai tolak ukur kecerdasan. Oleh karena itu, untuk membuktikan apakah hal tersebut benar atau tidak, akan dilakukan penelitian untuk mengetahui kecenderungan siswa SMP/MTs sederajat di Provinsi DI Yogyakarta, yang memiliki sebutan "Kota Pelajar" di mata masyarakat, dalam memilih sekolah lanjutannya.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, rumusan masalah dapat dituliskan sebagai berikut:

a. Apakah metode klasifikasi Decision Tree Algoritma C5.0 terbukti layak digunakan dalam data yang diuji?

- b. Bagaimana kecenderungan siswa SMP/MTs sederajat dalam memilih SMA atau SMK berdasarkan faktor-faktor yang telah ditentukan?
- c. Apa solusi yang dapat diambil dari hasil analisis yang telah dilakukan?

1.3 Tujuan Kerja Praktik

1.3.1 Tujuan Umum

- a. Terlaksananya mata kuliah wajib Kerja Praktik yang dilaksanakan pada waktu antar semester.
- b. Menjalin kerja sama dengan berbagai instansi/perusahaan.
- c. Memberikan pengalaman dan pengetahuan kepada mahasiswa dalam mengenal dunia kerja.

1.3.2 Tujuan Khusus

- a. Memberikan kesempatan kepada mahasiswa dalam mengenal dan merasakan dunia kerja.
- Melatih mahasiswa untuk solutif dan mandiri dalam menyelesaikan persoalan di dunia kerja.
- c. Mengetahui kemampuan mahasiswa dalam mengaplikasikan ilmu yang diperoleh selama perkuliahan.

1.4 Manfaat Kerja Praktik

1.4.1 Bagi Mahasiswa

- a. Mengenalkan dan memberikan pengalaman kerja kepada mahasiswa untuk mengaplikasikan pengetahuan yang diperoleh selama perkuliahan di tempat Kerja Praktik.
- b. Melatih kemampuan beradaptasi dan berkembang bagi mahasiswa dalam menghadapi berbagai situasi di dunia kerja.
- c. Meningkatkan rasa tanggung jawab, kedisiplinan, dan kemampuan mahasiswa dalam menangani situasi nyata di dunia kerja.

1.4.2 Bagi Program Studi Sarjana Statistika

- a. Menambah relasi instansi atau perusahaan tempat Kerja Praktik dengan Program Studi Sarjana Statistika.
- b. Memberikan kepercayaan bagi instansi atau perusahaan tempat Kerja Praktik dalam memberikan peluang kerja bagi mahasiswa Program Studi Statistika.

- c. Mengukur kemampuan dan pemahaman mahasiswa dalam mengaplikasikan pengetahuan yang telah diberikan dalam perkuliahan.
- d. Menjadi bahan tinjauan pustaka bagi pihak yang membutuhkan dalam meningkatkan kualitas mahasiswa Program Studi Statistika.

1.4.3 Bagi Instansi/Perusahaan

- a. Menjalin relasi dengan Program Studi Sarjana Statistika Universitas Gadjah Mada.
- Mengenalkan pekerjaan yang dilaksanakan oleh Dinas Pendidikan, Pemuda, dan Olahraga Provinsi DIY kepada mahasiswa.

1.5 Waktu dan Tempat Pelaksanaan

1.5.1 Waktu Pelaksanaan

Berdasarkan kalender akademik Universitas Gadjah Mada semester genap 2022/2023 serta kesepakatan dari (instansi), kerja praktik ini dilaksanakan selama 25 hari kerja dimulai dari tanggal 26 Juni hingga 28 Juli 2023.

1.5.2 Tempat Pelaksanaan

Kerja Praktik dilaksanakan di:

nama instansi : Dinas Pendidikan, Pemuda, dan Olahraga DIY

penempatan : Seksi Perencanaan dan Pengembangan Mutu

Pendidikan, Bidang Perencanaan

alamat instansi : Jalan Cendana 9 Yogyakarta, Yogyakarta 55166

nomor telepon instansi : (0274) 550330

1.6 Metodologi Penelitian

1.6.1 Metode Pengumpulan Data

Data pada penelitian ini menggunakan data sekunder yang bersumber dari Dinas Pendidikan, Pemuda, dan Olahraga (DIKPORA) Provinsi DIY.

1.6.2 Metode Analisis Data

Metode analisis data yang digunakan yaitu Decision Tree dengan menggunakan Algoritma C5.0 untuk menganalisis kecenderungan siswa SMP sederajat yang bersekolah di Provinsi Yogyakarta dalam memilih tingkatan akhir pendidikan menengah (SMA/SMK) pada Jalur Zonasi Reguler dari variabel jenis kelamin, tempat tinggal siswa (berdasarkan kabupaten/kota), serta 4 nilai ASPD siswa yang dimiliki.

BAB II

PROFIL DINAS PENDIDIKAN, PEMUDA, DAN OLAHRAGA DIY

2.1 Visi dan Misi

2.1.1 Visi

Birokrasi katalistik percepatan pembangunan pendidikan bangsa.

2.1.2 Misi

- 1. Menumbuhkan birokrasi pembelajaran, profesional, loyal, dan terukur.
- 2. Mengembangkan instansi budaya dan menghargai keberagaman.
- 3. Menciptakan instansi beretika dan menghargai keberagaman.

2.2 Tujuan dan Sasaran

2.2.1 Tujuan

Untuk mewujudkan visi dan misi pembangunan pendidikan DIY, dirumuskan tujuan strategis pembangunan pendidikan DIY sebagai berikut:

- 1. Mengembangkan pendidikan berkualitas yang merata untuk semua, berdaya saing, dan non diskriminatif.
- 2. Menghasilkan generasi muda berkarakter yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, cinta tanah air dan bangsa, berjiwa luhur, berbudaya, menjadi teladan, rela berkorban, kreatif, inovatif, serta profesional.
- 3. Mewujudkan peran DIY dalam menciptakan inovasi pendidikan.
- 4. Mewujudkan pendidikan yang responsif terhadap kebutuhan pembangunan.
- 5. Mewujudkan pemuda dan olahraga yang berkualitas, berdaya saing, dan berbudaya.
- 6. Meningkatkan layanan pendidikan, pemuda, dan olahraga yang akuntabel dan berbudaya.

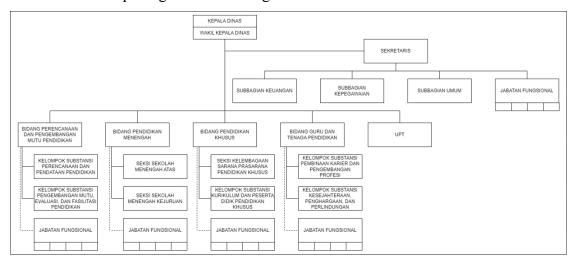
2.2.2 Sasaran

Adapun sasaran dari Dinas Pendidikan, Pemuda, dan Olahraga DIY yaitu terwujudnya seluruh tujuan yang telah ditetapkan.

2.3 Struktur Organisasi

Berdasarkan Peraturan Gubernur Daerah Istimewa Yogyakarta Nomor 68 Tahun 2018 tentang Kedudukan, Susunan Organisasi, Tugas, Fungsi, dan Tata Kerja Dinas

Pendidikan, Pemuda, dan Olahraga Bagian Kedua, bagan dari struktur organisasi instansi terkait dapat digambarkan sebagai berikut:



Bagan 1 Bagan Struktur Organisasi Dinas Pendidikan, Pemuda, dan Olahraga DIY

2.4 Tugas dan Fungsi

2.4.1 Tugas

Tugas dari Dinas Pendidikan, Pemuda, dan Olahraga DIY yaitu membantu Gubernur Provinsi DIY untuk melaksanakan urusan pemerintahan bidang pendidikan dan urusan pemerintahan kepemudaan dan olahraga.

2.4.2 Fungsi

Untuk melaksanakan tugasnya, Dinas Pendidikan, Pemuda, dan Olahraga DIY memiliki fungsi sebagai berikut

- Penyusunan program kerja Dinas Pendidikan, Pemuda, dan Olahraga DIY.
- Perumusan kebijakan teknis bidang pendidikan, pemuda, dan olahraga.
- Penyelenggaraan dan/atau pengelolaan pendidikan menengah dan pendidikan khusus.
- Penyediaan satuan pendidikan menengah dan pendidikan khusus.
- Fasilitasi dan pembinaan penyelenggaraan pendidikan layanan khusus.
- Penyediaan pendidik dan tenaga kependidikan, pendidikan menengah, dan pendidikan khusus.
- Penyediaan sarana dan prasarana pendidikan menengah dan pendidikan khusus.
- Pelaksanaan kurikulum muatan lokal pendidikan menengah dan pendidikan khusus.
- Penyusunan silabus dan buku teks pelajaran/modul pembelajaran pendidikan menengah dan pendidikan khusus.

BAB III

LANDASAN TEORI

3.1 Penerimaan Peserta Didik Baru (PPDB)

Penerimaan Peserta Didik Baru atau PPDB adalah penerimaan peserta didik baru pada Taman Kanak-Kanak (TK), Sekolah Dasar (SD), Sekolah Menengah Pertama (SMP), Sekolah Menengah Atas (SMA), dan Sekolah Menengah Kejuruan (SMK). PPDB untuk SD, SMP, dan SMA akan dilaksanakan melalui jalur pendaftaran PPDB. Jalur tersebut meliputi zonasi, afirmasi, perpindahan tugas orang tua/wali, dan/atau prestasi.

3.1.1 PPDB Jenjang Sekolah Menengah di Provinsi DIY

Penerimaan Peserta Didik Baru (PPDB) Jenjang Sekolah Menengah adalah kegiatan penerimaan calon peserta didik baru yang memenuhi syarat kelulusan jenjang SMP/MTs atau bentuk lain yang sederajat untuk memperoleh pendidikan dijenjang selanjutnya, baik Sekolah Menengah Atas (SMA) ataupun Sekolah Menengah Kejuruan (SMK). Dalam PPDB ini, terdapat maksimal jumlah peserta didik dalam Rombongan Belajar yang berjumlah 36 siswa. Rombongan Belajar adalah kelompok peserta didik yang terdaftar pada satuan kelas dalam satu sekolah. Pada Rombongan Belajar dalam PPDB SMA tidak dibagi berdasarkan jurusan, sesuai penerapan Kurikulum Merdeka, sedangkan pada Rombongan Belajar dalam PPDB SMK dibagi sesuai program studi/kompetensi keahlian. Kegiatan PPDB dilaksanakan oleh Sekolah SMA/SMK masing-masing sesuai dengan jadwal pelaksanaan PPDB. Pelaksanaan PPDB dapat menggunakan teknologi komunikasi dan informasi melalui sistem dalam jaringan (online) maupun sistem luar jaringan (offline).

3.1.1.1 Jalur Pendaftaran

Sama seperti PPDB pada umumnya, jalur pendaftaran PPDB untuk Sekolah Reguler meliputi zonasi, afirmasi, perpindahan tugas Orang Tua/Wali, dan prestasi. Namun, untuk Provinsi DI Yogyakarta, jalur zonasi terbagi menjadi dua, yaitu Zonasi Reguler dan Zonasi Radius. Zonasi Reguler adalah zonasi yang diukur berdasarkan jarak darat dari titik koordinat Kelurahan/Kalurahan/Desa ke SMA/SMK, sedangkan Zonasi Radius adalah wilayah dengan jarak tertentu yang mengelilingi titik koordinat SMA/SMK dengan mempertimbangkan kepadatan penduduk. Zonasi Reguler berdaya tampung 50% dari daya tampung sekolah, sedangkan Zonasi Radius berdaya tampung 5% dari daya tampung sekolah. Pada Zonasi Reguler, seleksi calon peserta didik baru

kelas X (sepuluh) dilakukan sesuai daya tampung berdasarkan ketentuan Rombongan Belajar dengan mempertimbangkan kriteria urutan prioritas, yaitu tempat tinggal calon peserta didik baru, nilai gabungan, pilihan sekolah dan/atau pilihan peminatan calon peserta didik, dan calon peserta didik baru yang mendaftar lebih awal. Pada Zonasi Radius, seleksi calon peserta didik baru dilakukan dengan kriteria urutan prioritas, yaitu jarak tempat tinggal sesuai domisili yang sah dengan titik koordinat sekolah, nilai gabungan, dan calon peserta didik baru yang mendaftar lebih awal. Nilai gabungan yang disebutkan sebelumnya adalah jumlah rata-rata nilai hasil perhitungan Rapor mata pelajaran Bahasa Indonesia, Matematika, Bahasa Inggris, dan IPA dari peserta didik SMP/MTs semester 1 sampai dengan semester 5 diberikan bobot 40%, ditambah jumlah nilai Asesmen Standarisasi Pendidikan Daerah (ASPD) diberikan bobot 55%, ditambah nilai Akreditasi Sekolah dikalikan 4 diberikan bobot 5%.

3.2 Decision Tree

Metode Decision Tree adalah metode penambangan data yang umum digunakan untuk mengklasifikasikan variabel berdasarkan beberapa kovariat atau untuk mengembangkan algoritma prediksi untuk variabel target. Metode ini mengklasifikasikan suatu populasi ke dalam segmen-segmen seperti cabang yang membentuk pohon terbalik dengan *root node*, *internal node*, dan *leaf node*. Algoritma ini bersifat non-parametrik dan dapat digunakan secara efisien untuk menangani kumpulan data besar dan rumit tanpa menerapkan struktur parametrik yang rumit. Jika ukuran sampel cukup besar, data penelitian dapat dibagi menjadi data pelatihan (*data train*), dan data validasi (*data test*). Data pelatihan tersebut digunakan untuk membuat model Decision Tree dan data validasi digunakan untuk menguji model yang telah didapat serta menentukan model akhir yang optimal.

3.2.1 Algoritma C5.0

Algoritma C5.0 merupakan perkembangan dari algoritma C4.5 yang mana juga perkembangan dari algoritma ID.3. Algoritma ID.3 (*Iterative Dichotomisher 3*) pertama kali ditemukan oleh John Ross Quinlan pada tahun 1986. Algoritma ini merupakan metode untuk membangun sebuat *Decision Tree* yang paling mendasar, di mana algoritmanya melakukan pencarian secara menyeluruh pada semua kemungkinan *Decision Tree*.

Kemudian, algoritma C4.5 merupakan perkembangan dari algoritma ID.3 di mana "C" menunjukkan algoritma ditulis dalam C dan versi spesifik algoritma 4.5. kriteria pemisahan yang digunakan C4.5 adalah pemerolehan informasi yang

dinormalisasi (perbedaan entropi). Atribut dengan informasi ternormalisasi tertinggi akan dipilih untuk mengambil keputusan. Algoritma C4.5 kemudian dilakukan berulang pada sub-list yang dipartisi. Algoritma ini telah mengalami peningkatan dari algoritma sebelumnya, seperti dapat menangani data kontinu maupun diskrit, menangani *missing values* pada *data train* dengan memberikan nilai berdasarkan nilai yang paling dominan, serta dapat memangkas banyak cabang dalam *decision tree* yang terbentuk atau *prunning* untuk memperkecil *decision tree* agar mudah dipahami. Namun, algoritma C4.5 masih memiliki kelemahan dalam menangani data berjumlah besar karena keterbatasan dalam menghitung entropi sehingga dapat menimbulkan *overlapping*.

Selanjutnya, sebagai perkembangan dari algoritma C4.5, dibuatlah algoritma C5.0. algoritma C5.0 telah menjadi standar untuk menghasilkan *decision tree* karena algoritma ini mampu menagani sebagai besar masalah secara langsung. Dibandingkan dengan metode lainnya, algoritma C5.0 umumnya memiliki kinerja yang hampir sama baiknya, tetapi lebih mudah dipahami dan diterapkan. Salah satu keuntungan dari algoritma ini adalah *prunning* yang dilakukan menangani banyak masalah secara otomatis menggunakan default yang cukup masuk akal. Strategi keseluruhannya adalah melakukan pemangkasan pasca pemangsaan pohon. Hal ini dilakukan terlebih dahulu pada *data train* yang mengalami *overfit*, kemudian *node* dan cabang yang memiliki pengaruh kecil terhadap kesalahan klasifikasi akan dihapus. Dibandingkan dengan algoritma C4.5, tingkat akurasi dari algoritma C5.0 lebih tinggi, pengambilan keputusan lebih cepat, dan penggunaan memori yang jauh lebih rendah.

3.3 Confusion Matrix

Confusion Matrix adalah matriks yang digunakan untuk mengukur performa model *machine learning* dari *data test*. Matriks ini sering digunakan untuk mengukur kinerja model klasifikasi, yang bertujuan untuk memprediksi label kategorik untuk setiap *input*. Pada matriks ini, akan ditampilkan jumlah *true positive* (TP), *true negative* (TN), *false positive* (FP), dan *false negative* (FN) yang dihasilkan oleh model pada *data train*.

Tabel 1 Tabel Confusion Matrix

Confusion Matrix	Data Seb	enarnya
	Ya	Tidak

Data Prediksi	Ya	True Positive (TP)	False Positive (FP)
	Tidak	False Negative (FN)	True Negative (TN)

Keterangan:

- TP: jika data sebenarnya bernilai 'ya' dan data prediksi bernilai 'ya'.
- FP: jika data sebenarnya bernilai 'tidak' dan data prediksi bernilai 'ya'.
- FN: jika data sebenarnya bernilai 'ya' dan data prediksi bernilai 'tidak'.
- TN: jika data sebenarnya bernilai 'tidak' dan data prediksi bernilai 'tidak'.

Dari hasil matriks yang telah didapat, peneliti dapat memperoleh

• Recall

Recall digunakan untuk mengukur efektivitas model klasifikasi dalam mengidentifikasi semua contoh yang relevan dari sekumpulan data.

$$Recall = \frac{TP}{TP + FN}$$

• Precision

Precision adalah ukuran yang digunakan untuk mengukur seberapa akurat prediksi positif pada sebuah model.

$$Precision = \frac{TP}{TP + FP}$$

• Accuracy

Accuracy adalah ukuran yang digunakan untuk mengukur kinerja dari suatu model.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Semakin besar nilai dari hasil pengukuran tersebut, akan semakin bagus model yang diperoleh. Adapun keakurasian nilai AUC dapat dikategorikan menjadi 5 kategori, antara lain:

- Akurasi bernilai 0.90 1.00 = Excelent Classification
- Akurasi bernilai $0.80 0.90 = Good\ Classification$
- Akurasi bernilai 0.70 0.80 = Fair Classification
- Akurasi bernilai 0.60 0.80 = Poor Classification
- Akurasi bernilai $0.50 0.60 = Failure\ Classification$

BAB IV

HASIL DAN PEMBAHASAN

4.1 Deskripsi Data

Data yang digunakan merupakan Data Penerimaan Peserta Didik Baru (PPDB) Jenjang Sekolah Menengah Jalur Zonasi Reguler Provinsi DI Yogyakarta 2023. Pertimbangan ini dilakukan karena ingin melihat kecenderungan siswa dalam memilih sekolah pilihan terhadap nilai ASPD yang diperoleh. Penggunaan nilai ASPD (dalam nilai gabungan) dinilai cenderung lebih dipertimbangkan oleh siswa yang mendaftar pada jalur zonasi reguler sehingga akan digunakan data PPDB pada jalur zonasi reguler. Variabel yang akan digunakan dalam analisis Decision Tree sebagai berikut:

- Jenis Kelamin Siswa
- Tempat Tinggal Siswa (Kota/Kabupaten)
- Nilai Ujian Asesmen Standarisasi Pendidikan Daerah (ASPD); Matematika, Bahasa Indonesia, Bahasa Inggris, dan Ilmu Pengetahuan Alam
- Pilihan Siswa; SMA atau SMK

Variabel Pilihan Siswa akan menjadi variabel dependen atau merupakan variabel uji dalam analisis ini.

4.2 Pre-Processing

Dengan menggunakan data PPDB, dilakukan pengecekan apakah data memiliki kolom yang kosong (null) dan diperoleh bahwa terdapat 8 baris yang memiliki kolom kosong (null) sehingga baris tersebut akan dihilangkan, didapatkan jumlah data yang akan dilakukan analisis sebanyak 14.985 data. Kemudian, oleh karena analisis Decision Tree lebih menganjurkan data yang bersifat kategorik, data PPDB yang memiliki variabel bernilai kontinu (nilai ujian ASPD) akan dilakukan transformasi data menjadi data kategorik dengan keterangan sebagai berikut:

Tabel 2 Tabel Transformasi Data

Variabel	Nilai
Jenis Kelamin	Laki-laki
	Perempuan
Tempat Tinggal Siswa	Bantul
	Gunung Kidul

	Kulon Progo
	Sleman
	Yogyakarta
Nilai Ujian ASPD;	≤ 55 (Low)
Matematika	56 – 70 (Medium)
	71 – 85 (High)
	85 – 100 (Very High)
Nilai Ujian ASPD; Bahasa	≤ 55 (Low)
Indonesia	56 – 70 (Medium)
	71 – 85 (High)
	85 – 100 (Very High)
Nilai Ujian ASPD; Bahasa	≤ 55 (Low)
Inggris	56 – 70 (Medium)
	71 – 85 (High)
	85 – 100 (Very High)
Nilai Ujian ASPD; Ilmu	≤ 55 (Low)
Pengetahuan Alam	56 – 70 (Medium)
	71 – 85 (High)
	85 – 100 (Very High)
Pilihan Siswa	SMA
	SMK

Dari data yang telah ditransformasikan, akan dilakukan *splitting* data dengan proporsi 70% (11.988 data) sebagai data train dan 30% (2.997 data) sebagai data test. Data train digunakan untuk membuat model Decision Tree, sedangkan data test digunakan untuk menguji model yang telah dibuat sebelumnya. Keakuratan hasil dari model Decision Tree tersebut akan dilihat dari nilai akurasi yang didapatkan dengan menggunakan Confusion Matrix.

4.3 Hasil Analisis

Analisis Decision Tree dilakukan dengan menghitung nilai Entropy dan Gain untuk menentukan node pertama atau *root node* pada Decision Tree yang akan dibuat.

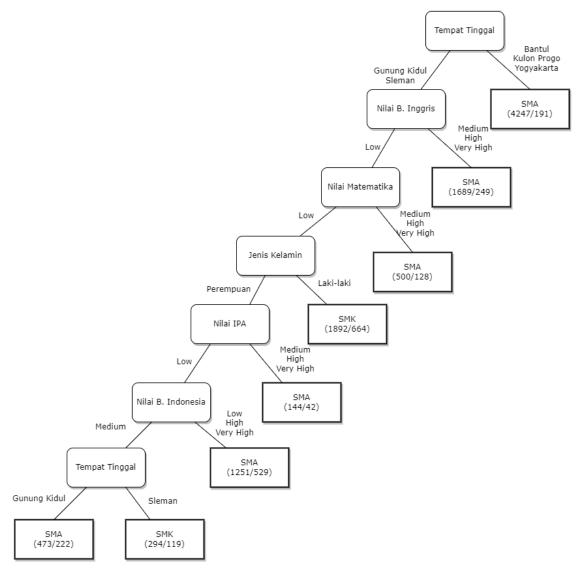
Tabel 3 Table Nilai Entropy dan Gain

Total	Atribut	Jumlah	SMA	SMK	Entropy	Gain
Total	7 tti ibut	11988	8830	3158	0.83188	
Jenis Kelamin	Laki-laki	5021	3245	1776	0.93734	0.02164
Jems Kelamin	Perempuan	6967	5585	1382	0.71864	
	Bantul	2539	2428	111	0.25909	0.14941
	Gunung Kidul	2988	1539	1449	0.99935	
Tempat Tinggal	Kulon Progo	1122	1100	22	0.13923	
	Sleman	4165	2682	1483	0.93937	
	Yogyakarta	1174	1081	93	0.39941	
	Low	8260	5457	2803	0.92419	0.05644
Nilai Ujian ASPD;	Medium	1600	1398	202	0.54706	
Matematika	High	1152	1051	101	0.42866	
	Very High	976	924	52	0.30016	
	Low	992	504	488	0.99981	0.07472
Nilai Ujian ASPD;	Medium	2702	1517	1185	0.98908	
Bahasa Indonesia	High	5026	3845	1181	0.78659	
	Very High	3268	2964	304	0.44648	
	Low	8004	5211	2793	0.93313	0.06413
Nilai Ujian ASPD;	Medium	2076	1842	234	0.50806	
Bahasa Inggris	High	1476	1360	116	0.39719	
	Very High	432	417	15	0.21755	
	Low	8049	5297	2752	0.92663	0.05492
Nilai Ujian ASPD;	Medium	2495	2184	311	0.54258	
Ilmu Pengetahuan Alam	High	1142	1059	83	0.37584	
	Very High	302	290	12	0.24108	

Nilai Gain terbesar akan digunakan sebagai node pertama dalam Decision Tree. Terlihat bahwa variabel Tempat Tinggal memiliki nilai Gain terbesar dibandingkan lainnya sehingga variabel tersebut digunakan sebagai node pertama.

4.3.1 Decision Tree Algoritma C5.0

Setelah dilakukan analisis Decision Tree dengan menggunakan RStudio dan diulang sebanyak 100 kali, diperoleh Decision Tree sebagai berikut.



Bagan 2 Bagan Hasil Decision Tree Algoritma C5.0

4.3.2 Confusion Matrix

Selanjutnya, dilakukan perhitungan Confusion Matrix dengan menggunakan RStudio, didapatkan sebagai berikut

Tabel 4 Tabel Hasil Confusion Matrix dari Decision Tree Algoritma C5.0

Confusion Matrix		Nilai Se	benarnya
Comusio	ii iviatrix	SMA	SMK
Nilai	SMA	2822	439
Prediksi	SMK	489	745

Kemudian, akan dilakukan perhitungan Accuracy, Precission, dan Recall dari Confusion Matrix yang telah didapat.

$$Recall = \frac{TP}{TP + FN} = \frac{2822}{2822 + 489} = 0.8523$$

$$Precision = \frac{TP}{TP + FP} = \frac{2822}{2822 + 439} = 0.8654$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{2822 + 745}{2822 + 745 + 439 + 489} = 0.7935$$

Didapatkan hasilnya sebagai berikut:

Accuracy	Precision	Recall
0.7935	0.8654	0.8523

Dapat dilihat bahwa nilai dari *Accuracy* dari model analisis ini adalah 0,7935 dimana nilai tersebut masuk ke dalam kategori *Fair Classification* pada kategori AUC. Hal ini membuat model Decision Tree Algoritma C5.0 yang dilakukan pada data ini memiliki akurasi yang tidak terlalu baik, tetapi tidak terlalu buruk juga. Namun, Decision Tree Algoritma C5.0 terbukti dapat mengatasi data yang berjumlah cukup besar dan menghasilkan hasil klasifikasi yang masih layak untuk digunakan.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan pada bab-bab sebelumnya, dapat disimpulkan diantaranya:

- Analisis Decision Tree C5.0 terbukti dapat mengklasifikasikan data PPDB Jalur Zonasi Reguler Sekolah Menengah Provinsi DI Yogyakarta dengan tingkat akurasi sebesar 0,7935 atau masuk ke dalam kategori fair classfication.
- Kecenderungan siswa dalam memilih SMA atau SMK untuk jenjang lanjutan setelah lulus dari SMP/MTs atau sederajat berdasarkan tempat tinggal siswa, yaitu Kabupaten/Kota siswa.
- Kabupaten Bantul, Kabupaten Kulon Progo, dan Kota Yogyakarta memiliki siswa yang mayoritas memilih SMA untuk jenjang lanjutan setelah lulus dari SMP/MTs atau sederajat.
- Siswa dari Kabupaten Gunung Kidul dan Kabupaten Sleman cenderung memperhatikan nilai ASPD yang diperoleh serta jenis kelamin, dengan urutan prioritas yaitu, Nilai Bahasa Inggris, Nilai Matematika, Jenis Kelamin, Nilai IPA, Nilai Bahasa Indonesia, kemudian kembali lagi pada tempat tinggal (Kabupaten Gunung Kidul dan Kabupaten Sleman).
- Kecenderungan siswa dalam memilih SMA/SMK berdasarkan nilai ASPD yang diperoleh yaitu siswa yang memiliki nilai High (70 – 85) dan Very High (85 – 100) akan memiliki SMA dibandingkan SMK.
- Siswa berjenis kelamin laki-laki cenderung memilih SMK apabila nilai Bahasa
 Inggris dan nilai Matematika mereka masuk dalam kategori Low (< 55).

5.2 Saran

Berdasarkan penelitian yang telah dilakukan, beberapa saran yang dapat diberikan untuk penelitian selanjutnya yaitu dapat terlihat bahwa kecenderungan dalam memilih SMA/SMK masih ada sehingga perlu dilakukan sosialisasi pada siswa SMP/MTs sederajat bahwa pemilihan SMA maupun SMK untuk jenjang lanjutan setelah lulus perlu dipikirkan secara matang berdasarkan tujuan siswa masing-masing, bukan berdasarkan tempat tinggal, nilai, maupun jenis kelamin. Kemudian, untuk penelitian selanjutnya dapat digunakan data

yang lebih lengkap (seperti kelengkapan data nilai gabungan atau faktor lainnya yang diperlukan) untuk menambah akurasi penelitian sehingga hasil yang diperoleh lebih baik.

DAFTAR PUSTAKA

- C5.0 Classification Models. (t.thn.). Dipetik Desember 10, 2023, dari cran-r project: https://cran.r-project.org/web/packages/C50/vignettes/C5.0.html
- CLASSIFICATION ALGORITHMS DECISION TREE. (t.thn.). Dipetik Desember 10, 2023, dari Intellegent Tutoring System for Programming: https://prutor.ai/classification-algorithms-decision-tree/
- Confusion Matrix in Machine Learning. (2023). Dipetik Desember 19, 2023, dari GeeksforGeeks: https://www.geeksforgeeks.org/confusion-matrix-machine-learning/
- Inayatus. (2019). *Prediksi Pasien Sakit Jantung Menggunakan Decision Tree*. Dipetik Desember 10, 2023, dari RPubs: https://rpubs.com/inayatus/decision-tree
- Indonesia. (2023). Peraturan Menteri Pendidikan dan Kebudayaan Nomor 1 Tahun 2021 tentang Penerimaan Peserta Didik Baru pada Taman Kanak-Kanak, Sekolah Dasar, Sekolah Menengah Pertama, Sekolah Menengah Atas, dan Sekolah Menengah Kejuruan.
- LU, Y.-y. S. (2015). Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry*, 27(2): 130–135. doi:https://doi.org/10.11919%2Fj.issn.1002-0829.215044
- Narkhede, S. (2018). *Understanding Confusion Matrix*. Dipetik Desember 19, 2023, dari Towards Data Science: https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62
- Oetama, N. B. (2020). Algoritma Klasifikasi Decision TreeC5.0 untuk Memprediksi Performa Akademik Siswa. *ejournals UMN*.
- Redaksi, T. (2016). *Mengenal Kantor Dinas Dikpora Provinsi Daerah Istimewa Yogyakarta*.

 Dipetik Desember 10, 2023, dari SMA NEGERI 1 LENDAH: http://sman1lendah.sch.id/2016/12/25/mengenal-kantor-dinas-dikpora-provinsi-d-i-yogyakarta/
- Sumitkrsharma. (2022). *Understanding C4.5 Decision tree algorithm*. Dipetik Desember 18, 2023, dari medium.com: https://medium.com/@sumit-kr-sharma/understanding-c4-5-decision-tree-algorithm-3bf0981faf4f
- Yobero, C. (2018). Determining Creditworthiness for Loan Applications Using C5.0 Decision Trees. Dipetik Desember 10, 2023, dari RPubs: https://rpubs.com/cyobero/C50

Yogyakarta, D. I. (2023). Peraturan Gubernur Nomor 15 Tahun 2023 tentang Petunjuk Teknis Penerimaan Peserta Didik Baru Sekolah Menengah Atas, Sekolah Menengah Kejuruan, dan Sekolah Luar Biasa.

LAMPIRAN

Lampiran 1 Surat Keterangan telah melaksanakan Kerja Praktek



PEMERINTAH DAERAH DAERAH ISTIMEWA YOGYAKARTA DINAS PENDIDIKAN, PEMUDA, DANOLAHRAGA

เมืองลง วิธีลเมืองลอกใหญ่เลา เลาเลืองกระบาม

Alamat :JalanCendana9 Yogyakarta55166 Telepon : (0274) 541322, 562278 Website : www.dikpora.jogiaprov.go.id | Email : dikpora@jogiaprov.go.id

SURAT KETERANGAN

NOMOR: 070/08853

Yang bertanda tangan di bawah ini :

Nama

: Drs. SUHIRMAN, M.Pd

NIP

: 19660317 199412 1 003

Pangkat/Golongan

: Pembina Utama Muda, IV/c

Jabatan

: Wakil Kepala Dinas Dinas Pendidikan, Pemuda, dan

Olahraga DIY

Instansi

: Dinas Pendidikan, Pemuda, dan Olahraga DIY

Denganinimenerangkanbahwa:

Nama

Silvy Aqila Maharani

NIS/NIM

478973

Sekolah / Universitas

Universitas Gadjah Mada Yogyakarta

Fakultas

Matematika dan IPA Statistika

Program Studi Tempat PKL

Seksi Perencanaan dan Pengembangan Mutu

Pendidikan Bidang Perencanaan

Nama tersebut di atas telah melaksanakan Magang/Kerja Praktek di Dinas Pendidikan, Pemuda, dan Olahraga Daerah Istimewa Yogyakarta selama 25 (dua puluh lima) hari kerja, terhitung mulai tanggal 26 Juni s.d. 28 Juli 2023

Selama melaksanakan Magang/Kerja Praktek di Dinas Pendidikan, Pemuda, dan Olahraga Daerah Istimewa Yogyakarta yang bersangkutan menunjukkan dedikasi yang baik, jujur dan penuh tanggungjawab terhadap tugas yang diberikan.

Surat keterangan ini dibuat untuk dipergunakan sebagaimana mestinya.

Yogyakarta, 2 8 JUL 2023

an. KEPALA

Drs. SUHTRMAN, M.Pd NIP., 19560317 199412 1 003

DISDI

Lampiran 2 Form Penilaian Kerja Praktek

Form Penilaian Kerja Praktek Program Studi Statistika FMIPA UGM

Nama NIM

: Silvy Aqila Maharani : 21/478973/PA/20777

Program Studi

: S1 Statistika

Nama Perguruan Tinggi

: Universitas Gadjah Mada

Tempat Kerja Praktik Waktu Kerja Praktik

: Dinas Pendidikan, Pemuda, dan Olahraga Provinsi DIY

: 26 Juni 2023 – 28 Juli 2023

No	Kriteria	Nilai (nilai dengan angka diantara 0 sampai 100)
1	Kedisiplinan	83
2	Pemahaman	81
3	Tanggung Jawab	81
4	Kreativitas	86
5	Kerjasama	8.5
6	Komunikasi	87
	Total Nilai	
	Rata-rata	

Catatan:

Petunjuk kisaran nilai huruf:

: $80 \le nilai \le 100$ A-: 75 ≤ nilai < 80 A/B : 70 ≤ nilai < 75 B+: $65 \le \text{nilai} \le 70$: 60 ≤ nilai < 65 : 50 ≤ nilai < 60 B-C: $35 \le \text{nilai} \le 50$ D : 25 ≤ nilai < 35 E : 0 ≤ nilai < 25

Yogyakarta,

(Andita Wahyu W