

# **Penerapan Algoritma Klasifikasi *Decission Tree* untuk Rekomendasi Jurusan Siswa Pada SMA N 1 Pati dengan Tes Bakat Minat**

**Dinda Dewani Khairunnisa Rahardyani**

**1900016158**

Skripsi diajukan kepada  
Fakultas Sains dan Teknologi Terapan  
sebagai bagian persyaratan untuk meraih derajat

**Sarjana Komputer**

pada Program Studi Sistem Informasi



Universitas Ahmad Dahlan  
Yogyakarta

Juli, 2023

**Skripsi**

**Penerapan Algoritma Klasifikasi *Decission Tree* untuk Rekomendasi Jurusan Siswa Pada SMA N 1 Pati dengan Tes Bakat Minat**

**Dinda Dewani Khairunnisa Rahardyani**

**1900016158**

Dipertahankan di depan Dewan Penguji

tanggal 29 Juli 2023



Iwan Tri Riyadi Yanto,  
S.Si., M.I.T., Ph.D.  
Ketua Pembimbing



Dr. Arif Rahman, S.Kom., M.T.  
Penguji 1



Farid Suryanto, S.Pd., M.T.  
Penguji 2

Diterima sebagai bagian  
persyaratan untuk meraih derajat  
Sarjana Komputer,



Dr. Yudi Ari Adi, S.Si., M.Si.  
Dekan Fakultas Sains dan  
Teknologi Terapan

# Pernyataan Tidak Plagiat

Saya yang bertanda tangan dibawah ini:

Nama : Dinda Dewani Khairunnisa Rahardyani  
NIM : 1900016158  
Email : dinda1900016158@webmail.uad.ac.id  
Program Studi : Sistem Informasi  
Fakultas : Fakultas Sains dan Teknologi  
Judul Tugas Akhir : Penerapan Algoritma Klasifikasi *Decision Tree* untuk  
Rekomendasi Jurusan Siswa Pada SMA N 1 Pati dengan  
Tes Bakat Minat

Dengan ini menyatakan bahwa:

1. Hasil karya yang saya serahkan ini adalah asli dan belum pernah mendapatkan gelar kesarjanaan baik di Universitas Ahmad Dahlan maupun di institusi pendidikan lainnya.
2. Hasil karya saya ini bukan saduran/terjemahan melainkan merupakan gagasan, rumusan, dan hasil pelaksanaan penelitian dan implementasi saya sendiri, tanpa bantuan pihak lain kecuali arahan pembimbing akademik dan narasumber penelitian.
3. Hasil karya saya ini merupakan hasil revisi terakhir setelah diujikan yang telah diketahui dan di setujui oleh pembimbing.
4. Dalam karya saya ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali yang digunakan sebagai acuan dalam naskah dengan menyebutkan nama pengarang dan dicantumkan dalam daftar pustaka.

Pernyataan ini saya buat dengan sesungguhnya. Apabila di kemudian hari terbukti ada penyimpangan dan ketidakbenaran dalam pernyataan ini maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya saya ini, serta sanksi lain yang sesuai dengan ketentuan yang berlaku di Universitas Ahmad Dahlan.

Yogyakarta, 8 Agustus 2023  
Yang Menyatakan



(Dinda Dewani Khairunnisa Rahardyani)

## PERNYATAAN PERSETUJUAN AKSES

Saya yang bertanda tangan di bawah ini:

Nama : Dinda Dewani Khairunnisa Rahardyani  
NIM : 1900016158  
Email : dinda1900016158@webmail.uad.ac.id  
Fakultas : Fakultas Sains dan Teknologi  
Program Studi : Sistem Informasi  
Judul Tugas Akhir : Penerapan Algoritma Klasifikasi *Decision Tree* untuk  
Rekomendasi Jurusan Siswa Pada SMA N 1 Pati dengan  
Tes Bakat Minat

Dengan ini saya menyerahkan hak *sepenuhnya* kepada Perpustakaan Universitas Ahmad Dahlan untuk menyimpan, mengatur akses serta melakukan pengelolaan terhadap karya saya ini dengan mengacu pada ketentuan akses tugas akhir elektronik sebagai berikut

Saya (**mengijinkan/tidak-mengijinkan**)\* karya tersebut diunggah ke dalam Repository Perpustakaan Universitas Ahmad Dahlan.

Demikian pernyataan ini saya buat dengan sebenarnya.

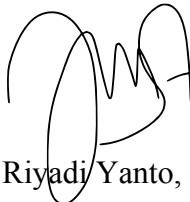
Yogyakarta, 8 Agustus 2023



Dinda Dewani Khairunnisa Rahardyani

Mengetahui,

Pembimbing\*\*



Iwan Tri Riyadi Yanto, S.Si., M.I.T., Ph.D.

### Ket:

\*coret salah satu

\*\*jika diijinkan TA dipublish maka ditandatangani dosen pembimbing dan mahasiswa

# Motto dan Persembahan

**“Ridho orang tua nomer satu, terutama ridho seorang ibu.”**

**“Thinking, feeling, action.”**

**“Jalani saja, nanti juga sampai. Berdoa saja, nanti juga dikabulkan.**

**Berusaha saja, nanti juga berhasil. Belajar saja, nanti juga paham. Tak perlu  
cemas soal masa depan, setiap langkah dan takdir terbaikmu sudah Allah  
perhitungkan.”**

Skripsi ini kupersembahkan untuk:

1. Allah SWT

Dengan ridho-Nya dan pertolongan-Nya, alhamdulillah skripsi saya jadi.

2. Keluarga tercinta

Terimakasih atas kepercayaan dan dukungan yang diberikan selama ini.

3. Armi, Kharisma, Nina, Wulan, dan Fika

Terimakasih sudah menjadi teman dan sahabat saya dalam dunia perkuliahan. Terimakasih sudah membantu saya jika dalam kesulitan.

Suka dan duka dilewati bersama. Semoga pertemanan kita till Jannah.  
Aamiin.

4. Teman – teman seperjuangan lainnya

Terimakasih atas dukungan dan motivasi.

5. Dosen Pembimbing

Terimakasih telah sabar membimbing saya dan selalu ada disaat saya bingung.

6. Para dosen Sistem Informasi UAD

# Kata Pengantar

*Bismillahirrahmaanirrahiim,*

*Assalamu'alaikum warahmatullahi wabarokatuh*

Segala puji bagi Allah swt., Tuhan semesta alam yang telah memberikan rahmat, hidayah, dan inayah kepada hambaNya, sehingga penulis dapat menyelesaikan Penelitian yang berjudul **Penerapan Algoritma Klasifikasi *Decission Tree* untuk Rekomendasi Jurusan Siswa Pada SMA N 1 Pati dengan Tes Bakat Minat.**

Penulis menyadari bahwa keberhasilan dalam menyelesaikan laporan ini berkat dorongan dan bimbingan berbagai pihak. Oleh karena itu, dalam kesempatan ini penulis menghaturkan terimakasih dan penghargaan kepada:

1. Bapak Dr.H. Yudi Ari Adi, S.Si., M.Si. selaku Dekan Fakultas Sains dan Teknologi Terapan Universitas Ahmad Dahlan.
2. Ibu Sri Handayaningsih, S.T., M.T. selaku Ketua Program Studi Sistem Informasi Universitas Ahmad Dahlan.
3. Bapak Iwan Tri Riyadi Yanto, S.Si., M.I.T. selaku Dosen Pembimbing.
4. Bapak Arif Rahman, S.Kom., M.T., selaku Dosen Penguji 1.
5. Bapak Farid Suryanto, S.Pd., M.T., selaku Dosen Penguji 2.
6. Seluruh dosen Sistem Informasi Universitas Ahmad Dahlan.

7. Guru – guru SMA N 1 Pati. Terutama guru BK, bagian TU, bagian Kurikulum, serta wali kelas yang sudah membantu saya mendapatkan data.

Selanjutnya penulis juga menyampaikan terimakasih kepada rekan-rekan satu perjuangan: Armi, Fika, Kharisma, Nina, Wulan, Salsa Alya, dkk.); dan berbagai pihak yang tidak dapat disebutkan satu persatu. *Jazakillah.*

Terakhir, penulis mengharapkan saran dan kritik yang membangun, karena skripsi ini sangat jauh dari sempurna. Semoga Allah meridhoi langkah kita dan skripsi ini dapat bermanfaat bagi siapapun yang membacanya.

*Aamiin.*

*Wassalamu 'alaikum warahmatullahi wabarokatuh*

Yogyakarta, 8 Agustus 2023



# Daftar Isi

|                                            | Halaman  |
|--------------------------------------------|----------|
| Halaman Judul.....                         | i        |
| Halaman Pengesahan .....                   | ii       |
| Halaman Pernyataan Tidak Plagiat .....     | iii      |
| Halaman Pernyataan Persetujuan Akses ..... | iv       |
| Halaman Motto dan Persembahan .....        | v        |
| Kata Pengantar .....                       | vii      |
| Daftar Isi.....                            | ix       |
| Daftar Gambar.....                         | xiv      |
| Daftar Tabel .....                         | xvii     |
| Daftar Kode Program .....                  | xix      |
| Daftar Algoritma .....                     | xx       |
| Abstrak.....                               | xxi      |
| <b>Bab 1. Pendahuluan .....</b>            | <b>1</b> |

|                                         |          |
|-----------------------------------------|----------|
| 1.1 Latar belakang .....                | 1        |
| 1.2 Identifikasi Masalah.....           | 4        |
| 1.3 Ruanglingkup.....                   | 4        |
| 1.4 Rumusan Masalah.....                | 5        |
| 1.5 Tujuan Penelitian .....             | 5        |
| 1.6 Manfaat Penelitian .....            | 5        |
| <b>Bab 2. Kajian Teori.....</b>         | <b>7</b> |
| 2.1 Kajian Penelitian Terdahulu .....   | 7        |
| 2.2 Kajian Teori .....                  | 10       |
| 2.2.1 <i>Data Mining</i> .....          | 10       |
| 2.2.2 Klasifikasi.....                  | 12       |
| 2.2.3 <i>Decission Tree</i> .....       | 14       |
| 2.2.4 Algoritma C4.5.....               | 15       |
| 2.2.5 Preprocessing Data .....          | 17       |
| 2.2.6 Evaluasi Klasifikasi.....         | 23       |
| 2.2.7 Tes Bakat Minat .....             | 26       |
| 2.2.7.1 Strength Typology (ST-30) ..... | 29       |

|                                                        |           |
|--------------------------------------------------------|-----------|
| <b>Bab 3. Metodologi .....</b>                         | <b>35</b> |
| 3.1 Subjek Penelitian .....                            | 35        |
| 3.2 Tahap Penelitian .....                             | 35        |
| 3.2.1 Pengumpulan Data .....                           | 36        |
| 3.2.2 Seleksi Data.....                                | 37        |
| 3.2.3 <i>Preprocessing</i> Data .....                  | 38        |
| 3.2.4 Transformasi Data .....                          | 38        |
| 3.2.4.1 <i>Label Encoding</i> .....                    | 39        |
| 3.2.5 <i>Splitting Data</i> .....                      | 39        |
| 3.2.6 Klasifikasi <i>Decision Tree</i> .....           | 39        |
| 3.2.7 Rekomendasi Jurusan.....                         | 40        |
| <b>Bab 4. Hasil dan Pembahasan.....</b>                | <b>41</b> |
| 4.1 Seleksi dan Pembersihan Data.....                  | 41        |
| 4.2 Transformasi Data .....                            | 44        |
| 4.3 Proses <i>Data Mining</i> .....                    | 47        |
| 4.3.1 Proses Klasifikasi Sebelum <i>Cleaning</i> ..... | 47        |
| 4.3.1.1 Klasifikasi .....                              | 47        |

|               |                                                  |           |
|---------------|--------------------------------------------------|-----------|
| 4.3.1.1.1     | <i>Splitting Data</i> .....                      | 47        |
| 4.3.1.1.2     | Evaluasi Klasifikasi .....                       | 51        |
| 4.3.2         | Proses Klasifikasi Setelah <i>Cleaning</i> ..... | 53        |
| 4.3.2.1       | Klastering Menggunakan <i>K-Prototypes</i> ..... | 53        |
| 4.3.2.1.1     | Hasil Klastering .....                           | 54        |
| 4.3.2.1.2     | Proses Cleansing Dengan Klastering .....         | 55        |
| 4.3.2.1.3     | Pengambilan Data Pure .....                      | 56        |
| 4.3.2.2       | Klasifikasi <i>Decision Tree</i> .....           | 57        |
| 4.3.2.2.1     | Evaluasi Klasifikasi .....                       | 58        |
| 4.4           | Perbandingan Hasil Dua Proses Klasifikasi .....  | 60        |
| 4.5           | <i>Streamlit</i> .....                           | 61        |
| 4.5.1         | Rekomendasi Jurusan .....                        | 62        |
| <b>Bab 5.</b> | <b>Penutup</b> .....                             | <b>64</b> |
| 5.1           | Kesimpulan .....                                 | 64        |
| 5.2           | Saran .....                                      | 65        |
| Daftar Acuan  | .....                                            | 66        |
| Lampiran 1    | .....                                            | 68        |

|                  |    |
|------------------|----|
| Lampiran 2 ..... | 69 |
| Lampiran 3 ..... | 70 |
| Lampiran 4 ..... | 77 |
| Lampiran 5 ..... | 84 |
| Lampiran 6 ..... | 85 |

# Daftar Gambar

|                                                                                      | Halaman |
|--------------------------------------------------------------------------------------|---------|
| Gambar 2.1 Data Mining sebagai bagian proses untuk menemukan pengetahuan (KDD) ..... | 11      |
| Gambar 2.2 Tahap membangun model .....                                               | 13      |
| Gambar 2.3 Tahap Menggunakan Model.....                                              | 13      |
| Gambar 2.4 Contoh <i>Decission Tree</i> .....                                        | 15      |
| Gambar 2.5 Tahapan Preprocessing.....                                                | 18      |
| Gambar 2.6 Contoh Binarization .....                                                 | 20      |
| Gambar 2.7 Contoh Mean Removal.....                                                  | 20      |
| Gambar 2.8 Contoh Scaling .....                                                      | 21      |
| Gambar 2.9 Contoh Normalization .....                                                | 22      |
| Gambar 2.10 Contoh Label Encoding.....                                               | 23      |
| Gambar 2.11 Model Confusion Matrix.....                                              | 23      |
| Gambar 2.12 Strength Map .....                                                       | 33      |

|                                                                                           |    |
|-------------------------------------------------------------------------------------------|----|
| Gambar 3.1 Tahap Penelitian.....                                                          | 36 |
| Gambar 4.1 Hasil Pengecekan Data.....                                                     | 44 |
| Gambar 4.2 <i>Library Decision Tree</i> .....                                             | 47 |
| Gambar 4.3 <i>Splitting Data</i> .....                                                    | 48 |
| Gambar 4.4 Proses Prediksi <i>Testing Data</i> .....                                      | 48 |
| Gambar 4.5 <i>Testing Data</i> .....                                                      | 49 |
| Gambar 4.6 Hasil <i>Confusion Matrix</i> dan Nilai Akurasi.....                           | 53 |
| Gambar 4.7 <i>Library K-Prototypes</i> .....                                              | 53 |
| Gambar 4.8 Output Proses Iterasi Prediksi Titik Centroid .....                            | 54 |
| Gambar 4.9 Titik <i>Centroid</i> yang Terbentuk.....                                      | 54 |
| Gambar 4.10 Proses <i>Looping</i> Memberi Label Sesuai Titik <i>Centroid</i> Terdekat.... | 55 |
| Gambar 4.11 Penambahan Kolom Hasil Klastering Pada Dataset .....                          | 55 |
| Gambar 4.12 Perhitungan <i>Confusion Matrix</i> Klastering.....                           | 56 |
| Gambar 4.13 Data <i>True Positive</i> dan <i>True Negative</i> .....                      | 56 |
| Gambar 4.14 Proses Prediksi <i>Testing Data</i> .....                                     | 57 |
| Gambar 4.15 <i>Testing Data</i> Setelah Cleaning.....                                     | 57 |
| Gambar 4.16 Hasil <i>Confusion Matrix</i> dan Nilai Akurasi Setelah Cleaning .....        | 60 |

|                                                 |    |
|-------------------------------------------------|----|
| Gambar 4.17 Simpan Model .....                  | 61 |
| Gambar 4.18 Tampilan Sistem.....                | 62 |
| Gambar 4.19 Hasil Rekomendasi Percobaan 1 ..... | 63 |



# Daftar Tabel

|                                                    | Halaman |
|----------------------------------------------------|---------|
| Tabel 2.1 Kajian Peneliti Terdahulu .....          | 9       |
| Tabel 2.2 Contoh Confusion Matrix .....            | 25      |
| Tabel 2.3 Tiga puluh Tipologi Manusia.....         | 29      |
| Tabel 4.1 Jumlah Data Siswa .....                  | 41      |
| Tabel 4.2 Sampel Data Asli Hasil Kuisisioner ..... | 41      |
| Tabel 4.3 Sampel Hasil Seleksi Data .....          | 43      |
| Tabel 4.4 Label Jenis Kelamin.....                 | 45      |
| Tabel 4.5 Label Asal SMP .....                     | 45      |
| Tabel 4.6 Label Tipologi.....                      | 45      |
| Tabel 4.7 Label Jurusan .....                      | 46      |
| Tabel 4.8 Sampel Hasil Pelabelan.....              | 46      |
| Tabel 4.9 Hasil Prediksi <i>Testing Data</i> ..... | 49      |

|                                                                     |    |
|---------------------------------------------------------------------|----|
| Tabel 4.10 <i>Confusion Matrix Data Test</i> .....                  | 52 |
| Tabel 4.11 Hasil Prediksi <i>Testing Data</i> dari Data Pure .....  | 58 |
| Tabel 4.12 <i>Confusion Matrix Data Test</i> Setelah Cleaning ..... | 59 |
| Tabel 4.13 Perbandingan Dua Proses Klasifikasi .....                | 60 |

# Daftar Kode Program

|                                                                                                    | Halaman |
|----------------------------------------------------------------------------------------------------|---------|
| Kode program 4.1. Mendeskripsikan kebutuhan klastering dan memprediksi titik <i>centroid</i> ..... | 54      |

# Daftar Algoritma

|                                                         |    |
|---------------------------------------------------------|----|
| Algoritma 4.1 Rule Decision Tree Sebelum Cleaning.....  | 51 |
| Algoritma 4.2 Rule Decision Tree Setelah Cleaning ..... | 58 |

# **Penerapan Algoritma Klasifikasi *Decision Tree* untuk Rekomendasi Jurusan Siswa Pada SMA N 1 Pati dengan Tes Bakat Minat**

## **Abstrak**

Salah satu SMA di kota Pati adalah SMA Negeri 1 Pati, yang memiliki dua jurusan: IPA dan IPS. Penjurusan merupakan salah satu upaya untuk siswa dalam memilih program pengajaran sesuai dengan kemampuan yang dianggapnya paling potensial. Tujuannya adalah mengelompokkan siswa berdasarkan minat, bakat, dan kecakapannya, serta untuk mempersiapkan melanjutkan studi dan dunia kerja. Pada tahun ajar 2022/2023 terjadi perubahan kurikulum yaitu menjadi kurikulum merdeka. Dimana pada kurikulum merdeka siswa dapat bebas memilih jurusan sesuai keinginan mereka. Karena hal tersebut maka tidak jarang siswa dan siswi kesulitan dalam memilih jurusan, sehingga mengambil penjurusan dengan asal-asalan dan hanya mengikuti teman. Hal tersebut menyebabkan ketidakcocokan antara penjurusan yang dipilih dengan potensi yang dimiliki. Hal ini dapat berdampak pada rendahnya prestasi belajar dan pemilihan jurusan pada perguruan tinggi.

Berdasarkan kasus tersebut pada penelitian ini menerapkan metode *Decision Tree* diharapkan dapat membantu siswa dalam merekomendasikan jurusan yang sesuai dengan bakat dan minat mereka. Tahapan yang dilakukan yaitu melakukan seleksi dan *preprocessing* data, transformasi data, proses klasifikasi dan pembuatan rekomendasi. Sedangkan atribut yang digunakan terdiri dari asal SMP, jenis kelamin, hasil tes bakat minat urutan 7 ke atas, nilai semester 1 (Matematika umum, Bahasa Indonesia dan Bahasa Inggris).

Hasil pengujian menggunakan *confusion matrix* pada proses klasifikasi sebelum cleaning diketahui tingkat akurasi dan presisi sebesar 63% dan 55% dengan 80% data *training* dan 20% data *testing*. Sedangkan hasil *confusion matrix* pada proses klasifikasi setelah cleaning terlebih dahulu diketahui tingkat akurasi dan presisi sebesar 94% dan 100% dengan besar *splitting data* sama.

Kata kunci: Tes Bakat Minat, Klasifikasi, Decision Tree, Rekomendasi Jurusan