



**PENERAPAN *DATA MINING* UNTUK PREDIKSI PENERIMAAN  
PESERTA DIDIK BARU SMP/MTS JALUR PMPA DI  
SMA NEGERI 8 PALEMBANG**

**SKRIPSI**

**KANIA AZKA AGUSTINE**

**161410135**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS BINA DARMA  
PALEMBANG  
2020**



**PENERAPAN *DATA MINING* UNTUK PREDIKSI PENERIMAAN  
PESERTA DIDIK BARU SMP/MTS JALUR PMPA DI  
SMA NEGERI 8 PALEMBANG**

**KANIA AZKA AGUSTINE**

**161410135**

**Skripsi ini diajukan sebagai syarat memperoleh gelar  
Sarjana Komputer**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS BINA DARMA  
PALEMBANG  
2020**

## **HALAMAN PENGESAHAN**

**PENERAPAN DATA MINING UNTUK PREDIKSI PENERIMAAN PESERTA  
DIDIK BARU SMP/MTS JALUR PMPA DI SMA NEGERI 8 PALEMBANG**

**KANIA AZKA AGUSTINE**

**161410135**

Telah diterima sebagai salah satu syarat untuk memperoleh gelar  
**Sarjana Komputer Program Studi Sistem Informasi**

Palembang, 29 Juni 2020

Fakultas Ilmu Komputer

Universitas Bina Darma

Pembimbing,



Andri, S.Kom., M.CS

Dekan,



The logo of Universitas Bina Darma Fakultas Ilmu Komputer is displayed, featuring the university's name in blue and red colors. A handwritten signature of "Dedy Syamsuar, S.Kom, M.I.T, Ph.D." is written across the logo.

Dedy Syamsuar, S.Kom, M.I.T, Ph.D.

## **HALAMAN PERSETUJUAN**

Skripsi berjudul "PENERAPAN DATA MINING UNTUK PREDIKSI PENERIMAAN PESERTA DIDIK BARU SMP/MTS JALUR PMPA DI SMA NEGERI 8 PALEMBANG" oleh KANIA AZKA AGUSTINE (161410135) telah dipertahankan didepan komisi pengaji pada hari SENIN tanggal 29 JUNI 2020.

### **Komisi Pengaji**

1. Ketua : Andri, S.Kom., M.CS.

2. Anggota : Susan Dian Purnamasari, M.Kom

3. Anggota : Maria Ulfa, M.Kom.

Mengetahui,

Program Studi Sistem Informasi

Universitas Bina Darma

Ketua,

  
Dr. Edi Surya Negara, M.Kom.

## **SURAT PERNYATAAN**

Saya yang bertanda tangan di bawah ini:

Nama : Kania Azka Agustine  
NIM : 161410135

Dengan ini menyatakan bahwa:

1. Karya tulis saya (Skripsi) adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (Sarjana) di Universitas Bina Darma atau perguruan tinggi lainnya;
2. Karya tulis ini murni gagasan, rumusan dan penelitian saya dengan arahan dari tim pembimbing;
3. Di dalam karya tulis ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dikutip dengan mencantumkan nama pengarang dan memasukkannya ke dalam daftar rujukan;
4. Saya bersedia tugas skripsi, jika cek keasliannya menggunakan *plagiarism checker* serta diunggah ke internet, sehingga dapat diakses secara daring;
5. Surat pernyataan ini saya tulis dengan sungguh-sungguh dan apabila terbukti melakukan penyimpangan atau ketidakbenaran dalam pernyataan ini maka saya bersedia menerima sanksi dengan peraturan dan perundang-undangan yang berlaku;

Demikian surat pernyataan ini saya buat agar dipergunakan sebagaimana mestinya.

Palembang, 29 Juni 2020

Yang membuat pernyataan.



Kania Azka Agustine

161410135

## **ABSTRAKSI**

Seleksi dalam Penerimaan Peserta Didik Baru (PPDB) melalui jalur Penelusuran Minat dan Prestasi Akademik (PMPA) adalah undangan bagi peserta didik untuk memasuki sekolah yang diinginkan tanpa melalui ujian tertulis. Dengan masalah ini, peneliti ingin mengetahui prediksi jalur PPDB SMP/MTs jalur PMPA berdasarkan nilai-nilai belajar peserta didik atau prestasi non akademik yang dimiliki peserta didik selama pendidikan SMP/MTs. Nilai belajar peserta didik saat belajar dan prestasi non-akademik di sekolah dapat dipercaya memiliki peran besar dalam menentukan tingkat kelulusan peserta didik tersebut.

Dengan memanfaatkan prestasi akademik dan prestasi non akademik peserta didik selama lima semester, mampu mendapatkan suatu informasi mengenai kelulusan peserta didik jalur PMPA menggunakan *data mining*. Adapun kategori yang dilihat dari kelulusan jalur PMPA melalui nilai peserta didik SMP/MTs, menggunakan metode klasifikasi *Decision Tree* (pohon keputusan) menggunakan algoritma C4.5. *Decision Tree* merupakan suatu metode mengubah data menjadi pohon keputusan dan aturan-aturan keputusan, dan algoritma C4.5 untuk membangun pohon keputusan dimulai dari pemilihan atribut sebagai akar, membuat cabang untuk tiap-tiap nilai, membagi kasus dalam cabang dan mengulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama dalam memprediksi kelulusan peserta didik jalur PMPA. Berdasarkan hasil analisis *data mining* menggunakan *RapidMiner* didapatkan hasil Total Nilai sebagai *root* (akar) dalam pembuatan decision tree dengan akurasi 87.84%.

**Kata Kunci:** *Data Mining*, Klasifikasi, *Decision Tree*, Algoritma C4.5, PMPA, Kelulusan.

## ABSTRACTION

Selection in the Admission of New Students (PPDB) through the Academic Interest and Achievement Tracking (PMPA) is an invitation for students to enter the desired school without going through a written exam. With this problem, the researcher wants to know the prediction of the PPDB Junior High School PMPA pathway based on the academic scores of students or the non-academic achievements of students during Junior High School education. The academic scores of students while studying and non-academic achievement at school can be trusted to have a big role in determining the pass rate of these students.

By utilizing the academic and non-academic achievements of students for five semesters, they are able to obtain information about the graduation of PMPA students using *data mining*. The categories seen from the graduation of the PMPA pathway through the scores of Junior High School students are using the classification method of *Decision Tree* (decision tree) and algorithm C4.5. *Decision Tree* is a method of converting data into a decision tree and decision rules, while the C4.5 algorithm for building a decision tree starts from selecting attributes as the root, creating branches for each value, dividing cases into branches and repeating the process for each branch until all cases in the branch have the same class in predicting graduation of PMPA pathway students. Based on the results of data mining analysis using *Rapid Miner*, the accuracy of total value results as *root* in making a decision is 89.19%.

**Keywords:** *Data Mining*, Classification, *Decision Tree*, C4.5 Algorithm, PMPA.

## **HALAMAN MOTTO DAN PERSEMBAHAN**

### **Motto:**

*MAN JADDA WA JADA*

“Siapa yang bersungguh-sungguh, dia akan berhasil” (Imam Syafii).

*LIKULLI SYAIIN THARIQ*

*WA TARIQUN JANNATIL ILMUN*

“Setiap sesuatu ada jalannya dan jalan ke surga adalah ilmu”

Alhamdulillah segala puji bagi Allah SWT yang telah memberikan rahmat dan karunia-Nya dalam penyusunan penelitian ini. Saya mempersembahkan tugas akhir skripsi ini dengan penuh rasa syukur dan terima kasih kepada:

1. Kedua orang tua tercinta, ayahanda Jurnaidi dan ibunda Yuliar Martina yang selalu memberikan dukungan, doa, dan kasih sayang yang tiada hentinya untuk anakmu. Terima kasih telah mengiringi setiap langkah ini dan mengajarkan banyak hal yang berharga serta memberikan yang terbaik. Harapannya kelak ilmu yang diperoleh ini bisa menjadi berkah dan membuat kalian bangga.
2. Ayuk tersayang, Debby Aulika Rachmah, yang selalu memberikan dukungan dan doa serta menjadi penyemangat dalam pembuatan penelitian ini. Terima kasih karena telah ikut meneman dan mewarnai perjalanan penelitian ini.
3. Sahabat-sahabatku yang setia mendampingi dan menyemangati di kala susah dan memberikan motivasi untuk terus maju. Terima kasih karena telah ikut meneman perjalanan penelitian ini.

## KATA PENGANTAR

Puji syukur kehadirat Allah SWT berkat rahmat dan hidayah-Nya penulis dapat menyelesaikan Laporan Praktek Kerja Lapangan yang berjudul **“Penerapan Data Mining Untuk Prediksi Penerimaan Peserta Didik Baru SMP/MTs Jalur PMPA Di SMA Negeri 8 Palembang”** tepat pada waktu yang ditentukan.

Dalam penulisan ini penulis menyadari bahwa tanpa adanya bimbingan, bantuan serta dukungan serta petunjuk dari semua pihak tidak mungkin laporan ini dapat diselesaikan. Oleh karena itu pada kesempatan ini penulis ingin menyampaikan rasa terima kasih kepada:

1. Allah SWT yang telah membimbing, membantu, dan menuntun saya selama pembuatan skripsi ini. Tanpa bimbingan dari-Nya, saya tidak akan pernah bisa mampu melewati semua rintangan selama proses pembuatan skripsi ini.
2. Orang tua dan Saudaraku tercinta yang telah memberikan bantuan baik moril maupun materil serta doa.
3. Dr. Sunda Ariana, M.Pd., M.M. selaku Rektor Universitas Bina Darma Palembang.
4. Dedy Syamsuar, S.Kom., M.IT., Ph.D., Selaku Dekan Fakultas Ilmu Komputer.
5. Dr. Edi Surya Negara, M.Kom., selaku Ketua Program Studi Sistem Informasi.
6. Andri, M.CS. selaku pembimbing yang telah membimbing peneliti dengan kesabaran, arahan, dan telah memberikan pelajaran yang sangat berharga.
7. Dosen Penguji I, Susan Dian Purnamasari, M.Kom dan Dosen Penguji II, Maria Ulfa, M.Kom yang telah bersedia menjadi penguji skripsi ini dan telah memberikan arahan yang sangat berharga.

8. Pihak SMA Negeri 8 Palembang telah menyetujui, mendukung dan membantu selama proses pembuatan skripsi.
9. Venny, Mifta, Intan, Indah, Ema, Aurel, Yeni, dan Annisa, yang telah memberikan motivasi, bantuan, serta dukungan, dan doa.
10. Teman seperjuanganku Sistem Informasi Angkatan 2016 yang telah yang telah memberikan kenangan dan pelajaran berharga kepada saya.
11. Semua pihak yang telah membantu dalam menyelesaikan skripsi ini yang tidak dapat disebutkan satu-persatu.

Akhir kata penulis menyadari sepenuhnya bahwa laporan ini sangat sederhana dan banyak kekurangan. Namun, harapan penulis kiranya skripsi ini dapat bermanfaat bagi kita semua.

Semoga Allah SWT melimpahkan taufik dan hidayah-Nya bagi kita semua. Amin.

**Palembang, Juni 2020**

**Peneliti**

## DAFTAR ISI

|   |             |
|---|-------------|
| <b>HALAMAN PENGESAHAN .....</b>             | <b>i</b>    |
| <b>HALAMAN PERSETUJUAN .....</b>            | <b>ii</b>   |
| <b>SURAT PERNYATAAN .....</b>               | <b>iii</b>  |
| <b>ABSTRAKSI .....</b>                      | <b>iv</b>   |
| <b>ABSTRACTION .....</b>                    | <b>v</b>    |
| <b>HALAMAN MOTTO DAN PERSEMBAHAN.....</b>   | <b>vi</b>   |
| <b>KATA PENGANTAR.....</b>                  | <b>vii</b>  |
| <b>DAFTAR ISI .....</b>                     | <b>ix</b>   |
| <b>DAFTAR GAMBAR.....</b>                   | <b>xii</b>  |
| <b>DAFTAR TABEL .....</b>                   | <b>xiii</b> |
| <b>BAB I PENDAHULUAN .....</b>              | <b>1</b>    |
| 1.1. Latar Belakang.....                    | 1           |
| 1.2. Perumusan Masalah .....                | 3           |
| 1.3. Tujuan Penelitian .....                | 3           |
| 1.4. Ruang Lingkup dan Batasan Masalah..... | 3           |
| 1.5. Manfaat Penelitian .....               | 4           |
| 1.6. Metodologi Penelitian.....             | 4           |
| 1.6.1. Waktu dan Tempat Penelitian.....     | 4           |
| 1.6.2. Alat .....                           | 5           |
| 1.7. Metode Penelitian .....                | 5           |
| 1.8. Metode Pengumpulan Data .....          | 5           |
| 1.9. Metode Analisis Data.....              | 6           |
| 1.10. Sistematika Penulisan.....            | 8           |
| <b>BAB II TINJAUAN PUSTAKA.....</b>         | <b>9</b>    |

|  |           |
|--|-----------|
| 2.1. Tinjauan Umum.....  | 9         |
| 2.1.1. Sejarah SMA Negeri 8 Palembang.....                     | 9         |
| 2.1.2. Visi dan Misi SMA Negeri 8 Palembang .....              | 10        |
| 2.1.3. Struktur Organisasi SMA Negeri 8 Palembang .....        | 11        |
| 2.2. Landasan Teori.....                                       | 12        |
| 2.2.1. <i>Data</i> .....                                       | 12        |
| 2.2.2. <i>Data Mining</i> .....                                | 12        |
| 2.2.3. Klasifikasi .....                                       | 13        |
| 2.2.4. <i>Decision Tree</i> (Pohon Keputusan).....             | 14        |
| 2.2.5. <i>Algoritma C4.5</i> .....                             | 15        |
| 2.2.6. <i>Pentaho Data Integration</i> .....                   | 16        |
| 2.2.7. <i>RapidMiner</i> .....                                 | 17        |
| 2.2.8. Evaluasi Hasil Klasifikasi dan Alat Ukur.....           | 18        |
| 2.2.9. Penelitian Terdahulu .....                              | 19        |
| <b>BAB III ANALISA DATA MINING .....</b>                       | <b>21</b> |
| 3.1. <i>Data Selection</i> .....                               | 21        |
| 3.2. <i>Preprocessing</i> .....                                | 26        |
| 3.2.1. <i>Data Cleaning</i> .....                              | 27        |
| 3.2.2. Integrasi Data.....                                     | 30        |
| 3.3. <i>Data Transformation</i> .....                          | 31        |
| 3.4. <i>Data Mining</i> Menggunakan <i>Decision Tree</i> ..... | 33        |
| 3.5. <i>Interpretation</i> .....                               | 44        |
| <b>BAB IV HASIL DAN PEMBAHASAN .....</b>                       | <b>46</b> |
| 4.1. Implementasi <i>RapidMiner</i> .....                      | 46        |
| 4.2. Pembahasan .....  | 49        |

|   |           |
|---|-----------|
| <b>BAB V PENUTUP .....</b>                      | <b>55</b> |
| 5.1. Kesimpulan .....                           | 55        |
| 5.2. Saran .....                                | 55        |
| <b>DAFTAR PUSTAKA .....</b>                     | <b>56</b> |
| <b>LAMPIRAN.....</b>                            | <b>59</b> |
| 1.    Data Mentah Penelitian.....               | 60        |
| a.    Sampel Data Mentah PMPA Tahun 2017 .....  | 60        |
| b.    Sampel Data Mentah PMPA Tahun 2018 .....  | 60        |
| c.    Sampel Data Mentah PMPA Tahun 2019 .....  | 61        |
| 2.    Hasil Data Penelitian .....               | 62        |
| 3.    Surat Izin Melaksanakan Penelitian.....   | 63        |
| 4.    Surat Keterangan Selesai Penelitian ..... | 64        |
| 5.    Plagiarisme .....                         | 65        |

## DAFTAR GAMBAR

|   |    |
|---|----|
| Gambar 2. 1. Struktur Organisasi SMA Negeri 8 Palembang .....                           | 11 |
| Gambar 2. 2. Struktur Susunan Pohon Keputusan .....                                     | 14 |
| Gambar 2. 3. Rumus <i>Entropy</i> .....   | 16 |
| Gambar 2. 4. Rumus <i>Gain</i> .....  | 16 |
| Gambar 2. 5. Rumus <i>Accuracy</i> .....  | 19 |
| Gambar 2. 6. Rumus <i>Precision</i> .....   | 19 |
| Gambar 2. 7. Rumus <i>Recall</i> .....  | 19 |
| Gambar 3. 1. Proses <i>Data Selection</i> .....   | 24 |
| Gambar 3. 2. Aplikasi <i>Pentaho Data Integration</i> untuk <i>Data Selection</i> ..... | 24 |
| Gambar 3. 3. Proses <i>Data Selection</i> di <i>Pentaho Data Integration</i> .....      | 25 |
| Gambar 3. 4. Proses Integrasi Data menggunakan <i>Pentaho Data Integration</i>          | 31 |
| Gambar 3. 5. Rumus Menghitung Entropi.....  | 35 |
| Gambar 3. 6. Rumus Menghitung Gain .....  | 35 |
| Gambar 3. 7. Pohon Keputusan <i>Node</i> Pertama.....                                   | 37 |
| Gambar 3. 8. Pohon Keputusan <i>Node</i> Kedua.....                                     | 40 |
| Gambar 3. 9. Pohon Keputusan <i>Node</i> Ketiga.....                                    | 42 |
| Gambar 3. 10. Pohon Keputusan <i>Leaf Akhir</i> .....                                   | 45 |
| Gambar 4. 1. Proses <i>Data Mining</i> menggunakan <i>Validation</i> .....              | 47 |
| Gambar 4. 2. Model Subproses <i>Validation Decision Tree</i> .....                      | 48 |
| Gambar 4. 3. <i>Accuracy Perfomance Decision Tree</i> .....                             | 49 |
| Gambar 4. 4. <i>Precision Perfomance Decision Tree</i> .....                            | 50 |
| Gambar 4. 5. <i>Recall Perfomance Decision Tree</i> .....                               | 50 |
| Gambar 4. 6. Pembentukan Pohon Keputusan.....   | 51 |
| Gambar 4. 7. Deskripsi dari <i>Decision Tree</i> .....                                  | 52 |
| Gambar 4. 8. Prediksi <i>Decision Tree</i> .....  | 53 |
| Gambar 4. 9. Hasil Prediksi yang Sesuai dari <i>Apply Model</i> .....                   | 53 |
| Gambar 4. 10. Hasil Prediksi yang Tidak Sesuai dari <i>Apply Model</i> .....            | 54 |

## **DAFTAR TABEL**

|   |    |
|---|----|
| Tabel 3. 1. Sampel Data yang Belum Diseleksi .....                    | 22 |
| Tabel 3. 2. Sampel Data yang Sudah Diseleksi.....                     | 25 |
| Tabel 3. 3. Format Sampel Data Sebelum <i>Preprocessing</i> .....     | 28 |
| Tabel 3. 4. Format Sampel Data Setelah <i>Preprocessing</i> .....     | 29 |
| Tabel 3. 5. Perubahan dalam <i>Data Transformation</i> .....          | 32 |
| Tabel 3. 6. Data Sampel Kasus SMA Negeri 8 Palembang .....            | 33 |
| Tabel 3. 7. Data Perhitungan <i>Algoritma C4.5 Node</i> Pertama.....  | 35 |
| Tabel 3. 8. TotalNilai $\leq 1743$ .....                              | 37 |
| Tabel 3. 9. Data Perhitungan <i>Algoritma C4.5 Node</i> Kedua.....    | 38 |
| Tabel 3. 10. Rank Sem 1, $>1$ .....                                   | 40 |
| Tabel 3. 11. Data Perhitungan <i>Algoritma C4.5 Node</i> Kedua .....  | 41 |
| Tabel 3. 12. Rank Sem 4, $>2$ .....                                   | 42 |
| Tabel 3. 13. Data Perhitungan <i>Algoritma C4.5 Node</i> Ketiga ..... | 43 |