

**Penerapan Algoritma C4.5 untuk Seleksi Penerimaan Siswa Baru pada SD
Islam Terpadu Permata Bunda Demak**
**Implementation of Decision Tree Algorithm for Selection of New Student
Admission on Permata Bunda Integrated Islamic Elementary School**

Linda Monizah Fitriani¹

Andik Setyono²

^{1,2}Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro
Jl. Nakula 1, No. 5-11 Semarang, Kode Pos 50131, Telp. (024) 3515261, Fax: 3569684
e-mail: 1, andik.setyono@dsn.dinus.ac.id

Diterima: 3 Juli 2018/ Disetujui : 19 Juli 2018

ABSTRACT

Proses seleksi penerimaan siswa baru merupakan hal mendasar yang dapat menentukan kelancaran kegiatan belajar mengajar di sekolah. Proses tersebut membutuhkan ketelitian agar hasilnya tepat dan akurat. Proses seleksi siswa baru terbagi menjadi 2 jenis penyeleksian, yaitu tes penjurusan dan tes wawancara. Tujuan dari penelitian ini adalah untuk membantu pihak sekolah dalam melakukan seleksi calon siswa sehingga dapat menjadi pendukung keputusan penerimaan siswa baru. Maka, diperlukannya pendekatan data mining agar menghasilkan informasi yang dapat mendukung pengambilan keputusan dalam penerimaan siswa baru. Algoritma yang digunakan adalah decision tree C4.5. Algoritma C4.5 dapat mendukung pengambilan keputusan penerimaan siswa baru melalui rules yang dihasilkan. Proses pengujian dengan RapidMiner menghasilkan akurasi sebesar 90,50%. Berdasarkan pengujian tersebut, peneliti mengolah kembali ke dalam bentuk aplikasi untuk membantu pihak sekolah. Maka, dilakukan penyebaran kuisisioner kepada pihak sekolah untuk menguji peranan aplikasi tersebut yang berupa 10 pertanyaan oleh 20 guru dan mendapatkan hasil index sebesar 81,5%. Dengan demikian, pihak sekolah puas terhadap aplikasi tersebut dan dapat membantu proses seleksi oleh pihak sekolah.

Kata kunci : *Seleksi, Penerimaan, Siswa, Data Mining, Klasifikasi, Decision Tree, C4.5*

ABSTRAK

The new admission process of selection is a basic rules for determining studying and learning in schools. This process requires precision so that the results are accurate and precise. The selection process for new students are divided into two types of screening, assessment tests and interviews. The purpose of this study is to assist schools in selecting prospective students so that they can be a decision support for new students. Thus, the need for data mining approach to generate information that can support decision-making for new admissions. The algorithm used is a C4.5 decision tree. C4.5 algorithms can support decision making new admissions through the rules generated. The testing process with RapidMiner yield 90.50% accuracy. Based on these tests, the researchers reprocess into the application form to help the school. So, do the questionnaire to the school to investigate the role of applications in the form of 10 questions by 20 teachers and an index of 81.5%. Thus, schools are satisfied with the application and can help the selection process by the school.

Keywords : *Selection, Admission, Students, Data Mining, Classification, Decision Tree, C4.5*

1. PENDAHULUAN

Pelaksanaan penerimaan siswa baru merupakan hal terpenting bagi sekolah. Hal ini dikarenakan penerimaan siswa baru merupakan titik awal yang dapat menentukan kelancaran kegiatan belajar mengajar di sekolah. Maka, proses penerimaan siswa baru harus sudah selesai dilaksanakan menjelang tahun ajaran baru [[HYPERLINK \l "Suh02" 1](#)]. Menurut Hamalik, peserta didik didefinisikan sebagai unsur masukan dalam instansi pendidikan, kemudian akan dimasukkan pada suatu proses pendidikan. Dengan demikian, tujuan pendidikan nasional dapat terwujud, yaitu terciptanya pribadi yang berkualitas [2]. Hal-hal yang dibutuhkan dalam penerimaan siswa baru seperti: (a) menentukan atau menyusun persyaratan pendaftaran untuk murid baru, (b) menyediakan formulir untuk pendaftar, (c) memberikan pengumuman, (d) menyediakan buku rekapan untuk pendaftar, (e) menentukan kuota siswa baru yang akan diterima [3].

Setiap tahunnya, SD Islam Terpadu Permata Bunda Demak melakukan penerimaan siswa baru. Pada penerimaan siswa baru, pihak sekolah perlu melakukan sebuah dokumentasi pendaftaran dan penyeleksian. Sampai saat ini, pihak sekolah mengolah data seleksi calon siswa baru dengan menggunakan *Microsoft Excel*. Cara tersebut membuat pihak sekolah merasa kurang puas, lantaran pengerjaan melakukan seleksi calon siswa baru membutuhkan waktu yang lama. Penyeleksian siswa baru membutuhkan ketelitian agar hasilnya tepat dan akurat. Karena, dalam proses seleksi siswa baru di sekolah tersebut terbagi menjadi 2 jenis penyeleksian, yaitu tes peninjauan dan tes wawancara.

Dalam proses seleksi penerimaan siswa baru, perlu melakukan perubahan metode penerimaan. Sehingga, proses tersebut memerlukan metode untuk menyelesaikan masalah di atas supaya menghasilkan informasi agar dapat dijadikan pendukung pengambilan keputusan dalam penerimaan siswa baru secara cepat dan akurat. Dengan demikian, hal yang perlu dilakukan yaitu melakukan pendekatan data mining untuk menentukan siswa baru.

Pendekatan data mining akan menghasilkan pendukung keputusan siswa baru yang lebih akurat dan tidak memakan waktu lama. Pada kasus ini, penggunaan teknik klasifikasi merupakan solusi yang tepat lantaran teknik tersebut dapat dijadikan informasi dalam mendukung pengambilan keputusan penerimaan siswa baru. Algoritma yang digunakan dalam klasifikasi salah satunya adalah decision tree C4.5.

Pada penelitian sebelumnya, Fitroh (2015) memiliki peningkatan nilai akurasi, yaitu dari 91,9% menjadi 94,6% [4]. Kemudian, Castaka Agus Sugiarto (2015) juga melakukan penelitian untuk menentukan Hasil Seleksi Masuk SMAN 1 Gibeber. Dalam penelitian tersebut, pengujian menggunakan Algoritma C4.5 menghasilkan Accuracy 99.05%, sedangkan dengan Neural Network menghasilkan Accuracy 95.02%. Algoritma Naive Bayes menghasilkan Accuracy 90.62%. Dari data tersebut diperoleh akurasi tertinggi pada penggunaan Algoritma C4.5 [5]. Kemudian penelitian dari Melda Khusmawathy menggunakan algoritma *Quantitative association rule* untuk menyeleksi calon siswa baru. Hasilnya, dari semua hasil tes setiap calon siswa harus lulus pada tes Psikotes [6].

Dengan demikian, Penerapan algoritma C4.5 diharapkan dapat menentukan siswa baru yang diterima secara akurat dan dapat mempersingkat waktu, dengan berbagai kriteria yang telah disebutkan diatas.

2. METODE PENELITIAN

2. 1 Prosedur Pengumpulan Data

Dataset yang digunakan diperoleh berdasarkan data primer yaitu sumber data yang didapat langsung dari SD Islam Terpadu Permata Bunda berupa data seleksi calon siswa baru berupa bilangan, yaitu skor penilaian baca, tulis, hitung, dan wawancara. Berikut ini merupakan contoh dataset seleksi siswa baru di SD Islam Terpadu Permata Bunda Demak :

Tabel 1. Contoh Dataset Seleksi Siswa Baru

No Pendaftar	Tahun	Nama	Asal Tk	Penjajagan			Jumlah Penjajagan	Wawancara					Jumlah Wawancara	Total Score	Status
				Baca	Tulis	Hitung		Konsepsio mal	Kematang an	Kondisi Rumah	Operaslon al	Kesehatan			
001	2015	Muhammad Zakii Marisk	Tk Nabila	10	5	10	25	3	2	2	2	3	12	37	Tidak Diterima
002	2015	Haikal Aminudin Faqih	Paud Assyfa	10	15	25	50	3	2	3	3	3	14	64	Diterima
003	2015	Mufti Deka Avrielano Az-Zuhri	Tkit Permata Bunda	25	20	25	70	3	3	3	3	3	15	85	Diterima
004	2015	Asfa Laenal Hawa	Tkit Permata Bunda	15	10	25	50	2	2	2	2	3	11	61	Diterima
005	2015	Fachrizal Putrandita Satriya Pratama	Tkit Permata Bunda	25	23	25	73	3	3	3	3	3	15	88	Diterima
006	2015	Jenni Nabila Srilarissa	Tk Aba 2 Pucang Gading	24	18	18	60	3	3	3	3	3	15	75	Diterima
007	2015	Dwi Rizky Suchi Amelia	Tk Al Ikhlas	25	25	25	75	3	2	2	3	2	12	87	Diterima
008	2015	Fadhil Teguh Wicaksono	Tkit Permata Bunda	25	21	25	71	3	3	3	3	3	15	86	Diterima
009	2015	Ariq Akhdan Wiratama	Paud Terpadu An-Nahl	20	15	15	50	3	2	3	3	3	14	64	Diterima
010	2015	Rayhan Fatan Assakhi	Tk Pembina	5	10	25	40	3	3	3	3	3	15	55	Tidak Diterima

2. 2 Desain Penelitian

Pada penelitian ini, penulis menerapkan model standarisasi data *Cross Industry Standart Process for Data Mining* (CRISP-DM), yaitu [7]:

1. *Business Understanding* atau Pemahaman Bisnis

Penelitian tersebut mampu mengolah sebuah data yang didapat dari SD Permata Bunda menjadi formula atau rules melalui proses data mining dengan cara menyiapkan strategi awal hingga metode yang digunakan untuk mencapai tujuan bisnis.

2. *Data Understanding* atau Pemahaman Data

Pada tahun ajaran 2014/2015, data seleksi siswa berjumlah 131 data. Sedangkan di tahun ajaran 2015/2016, data seleksi calon siswa berjumlah 128 data. Sedangkan di tahun 2016/2017, terdapat

121 data. Sehingga, total data mentah berjumlah 380 data. Data seleksi calon siswa baru tersebut terdiri dari atribut nomor pendaftaran, nama siswa, asal sekolah, skor tes baca, skor tes tulis, skor tes hitung, total skor wawancara, dan status.

3. *Data Preparation* atau Pengolahan Data

Adapun penjelasan dalam tahapan pengolahan data adalah sebagai berikut:

- Data cleaning and data selection*, yaitu tahapan awal dalam proses data mining. Tahap ini dilakukan untuk membuang atribut yang kurang lengkap atau yang tidak dibutuhkan agar lebih efisien, seperti jumlah dan keterangan.
- Data integration and data transformation*, yaitu tahap memindahkan database dan memproses transformasi data. Pada data seleksi calon siswa baru, data dari file yang berbeda seperti data pada tahun ajaran 2015/2016 dan 2016/2017 dipindah menjadi satu *file*. Kemudian dilakukan transformasi data dengan mengubah bentuk numerik menjadi nominal dalam bentuk interval.
- Data Reduction*, yaitu penggunaan data sesuai atribut dan jumlah *record* yang dibutuhkan. Jadi, beberapa *field* perlu dihilangkan karena tidak lengkapnya data. Terdapat 9 atribut yang nantinya akan diolah.

4. *Modelling* atau Pemodelan

Diawali dengan pemrosesan dataset seperti konversi data dan *replacing data*, kemudian dilakukan pemodelan menggunakan algoritma C4.5. Setelah proses pemodelan selesai, dilakukan proses evaluasi dengan menggunakan *Confusion Matrix*.

5. *Evaluation* atau Evaluasi

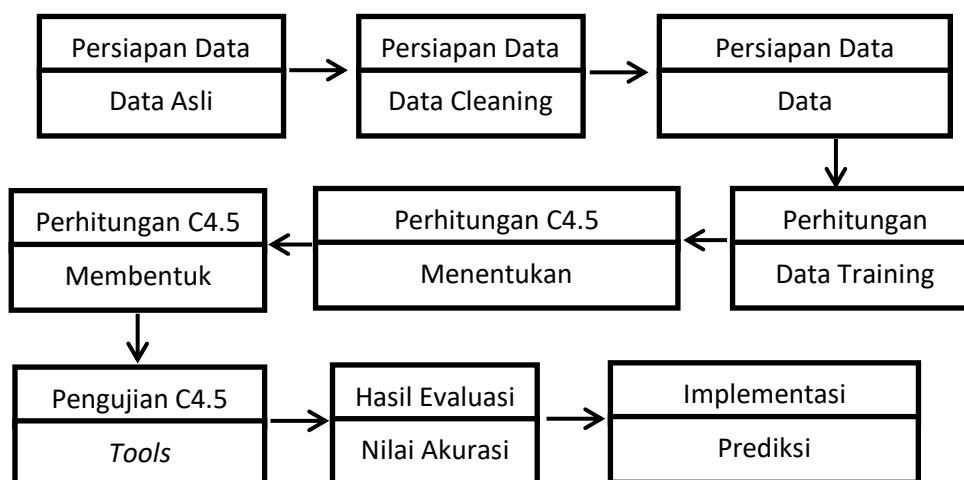
Pada langkah ini, penulis melakukan validasi, yaitu mengukur akurasi dari sebuah pemodelan algoritma yang akan diterapkan oleh *tools* RapidMiner. Pengujian yang dilakukan yaitu dengan *Confusion Matrix* untuk mengukur tingkat akurasi tersebut.

6. *Deployment* atau Penyebaran

Pada tahap ini, penelitian yang telah dilakukan akan menghasilkan analisis sehingga terbentuk *Decision Support System* (DSS) yang menghasilkan suatu produk untuk pihak sekolah, khususnya panitia penerimaan siswa baru sebagai bahan pertimbangan dalam menentukan siswa yang diterima di SD Islam Terpadu Permata Bunda.

2.3 Metode yang Diusulkan

Di bawah ini merupakan bagan mengenai langkah- langkah metode yang diterapkan secara rinci, yaitu :



Gambar 1