

Klasifikasi Peserta Didik Menggunakan Algoritma C4.5

Firda Amalia¹; Amanda²; Muhamad Aditya Purnama³

^{1,2,3}Program Studi Teknologi Rekayasa Perangkat Lunak

Politeknik Kelapa Sawit Citra Widya Edukasi – Bekasi

Email Penulis Korespondensi: 1frdalydrs@gmail.com

Abstrak

Pendidikan merupakan hal penting dalam meningkatkan kualitas sumber daya manusia. Dalam proses pendidikan, hal yang menjadi perhatian adalah pemetaan atau klasifikasi peserta didik. Peserta didik adalah makhluk individu dengan kepribadian yang memiliki sifat unik yang cocok untuk pertumbuhan dan perkembangan (Anggraeni, et al., 2022.). Data yang diperoleh dari operator TK Sarinah Bogor merupakan data yang valid dan mencakup jumlah peserta didik yang terdaftar dalam kelas A dan kelas B selama 11 tahun ajaran (2012/2013 hingga 2022/2023). Algoritma C4.5 dengan implementasi menggunakan Orange dapat menghasilkan model klasifikasi yang cukup baik dengan decision tree yang terdiri dari 15 nodes dan 8 leaves. Model ini memberikan tingkat Accuracy (keakuratan) sebesar 72.5% serta Area Under Curve (AUC) sebesar 0.737, serta nilai lain seperti F1 Score, Precision, Recall, dan Matthews Correlation Coefficient (MCC) menunjukkan kinerja model yang memadai.

Kata Kunci

Klasifikasi peserta didik, Algoritma C4.5, Pendidikan, Orange, Decision Tree, VBA.

Abstract

Education is an important factor in improving the quality of human resources. In the educational process, one of the focal points is the mapping or classification of students. Students are individual beings with unique personalities suitable for growth and development (Anggraeni, et al., 2022). The data obtained from the Sarinah Bogor Kindergarten operator is valid and covers the number of students registered in Class A and Class B for 11 academic years (2012/2013 to 2022/2023). The C4.5 algorithm, implemented using Orange, can generate a reasonably good classification model with a decision tree consisting of 15 nodes and 8 leaves. This model provides an Accuracy rate of 72.5% and an Area Under the Curve (AUC) of 0.737, as well as other values such as F1 Score, Precision, Recall, and Matthews Correlation Coefficient (MCC) indicating satisfactory model performance.

Keywords

Student Classification, C4.5 Algorithm, Education, Orange, Decision Tree, VBA.

Pendahuluan

Dalam konteks pendidikan di Indonesia, klasifikasi peserta didik sering kali dilakukan dengan cara yang konvensional, yaitu dengan cara melihat nilai rapor, tes potensi akademik (TPA), prestasi dibidang ekstrakurikuler, kedisiplinan dan kehadiran. Namun, cara tersebut cenderung subyektif dikarenakan metode tersebut bergantung pada beberapa faktor yang dapat mempengaruhi hasil evaluasi peserta didik.

Faktor yang mempengaruhi antara lain motivasi dan minat belajar, dukungan dan bimbingan, serta gaya belajar. Nilai rapor, meskipun memberikan gambaran tentang prestasi akademik, tidak mencerminkan potensi sebenarnya dari seorang siswa. Selain itu, tes potensi akademik (TPA) juga memiliki keterbatasan dalam mengukur kemampuan siswa secara menyeluruh.

Penelitian ini dilakukan dengan mengambil data dari peserta didik yang bersekolah di TK Sarinah Bogor selama periode tahun ajaran 2012 hingga 2022. Metode yang digunakan dalam penelitian ini adalah algoritma C4.5 untuk membangun model klasifikasi. Selain itu, pengujian data penelitian dilakukan menggunakan perangkat lunak *Orange* dan *VBA*.

Tujuan utama dari penelitian ini adalah untuk menguji dan menerapkan algoritma C4.5 dalam konteks klasifikasi peserta didik di TK Sarinah Bogor. Dengan algoritma ini, kita akan dapat mengkategorikan peserta didik berdasarkan berbagai faktor yang relevan. Selain itu, penelitian ini juga bertujuan untuk membangun sebuah model klasifikasi yang efisien dan akurat. Model ini akan menjadi alat yang berharga dalam pengelolaan data peserta didik di TK Sarinah Bogor, membantu pengambilan keputusan yang lebih baik, serta meningkatkan efisiensi dalam pengelolaan informasi peserta didik tersebut. Dengan demikian, tujuan penelitian ini adalah untuk memberikan kontribusi yang positif dalam meningkatkan manajemen pendidikan di TK Sarinah Bogor.

Metodologi

Teknik Pengumpulan Data

Teknik pengumpulan data merupakan metode atau cara yang digunakan untuk mengumpulkan informasi dan fakta yang relevan dalam sebuah penelitian atau studi. Tujuan dari teknik pengumpulan data untuk mendapatkan data yang akurat, dapat diandalkan, dan sesuai dengan pertanyaan penelitian atau tujuan yang ingin dicapai.

Teknik Pengolahan Data

Serangkaian metode dan prosedur yang digunakan untuk mengumpulkan, membersihkan, menganalisis, menginterpretasi, dan menyajikan data agar dapat digunakan untuk mengambil keputusan atau mendapatkan wawasan yang berharga. Pengolahan data melibatkan transformasi data mentah menjadi bentuk yang lebih berguna, relevan, dan dapat dimengerti.

Tabel 1 Data Peserta Didik TK Sarinah Bogor Tahun Ajaran 2012/2013

No.	Nama	Tempat, Tanggal Lahir	Alamat
1	Dwi Novia Syafitri	Depok, 3 November 2006	Buana Asri Blok A2 No. 2
2	Muhammad Adrian Eka S.	Cibinong, 9 Agustus 2007	Buana Asri Blok E3 No. 20
3	Kaniya Rahmah Ilahi	Bogor, 1 November 2007	Bogor Asri Blok A2 No. 8
4	Maria Aulia Sianturi	Depok, 2 Agustus 2008	Bogor Asri Kopasus A2 No. 14
5	Moreno Septianto	Bogor, 3 September 2007	Permata Bintang Blok F No. 11
6	Dedy Maulana Iskandar	Ngawi, 23 Januari 2007	Buana Asri Blok A1 No. 11
7	Sri Wulan Dari	Bogor, 10 September 2006	Kp. Tarikolot, RT02/RW06
8	Oliv Putri Vianti	Banyumas, 9 Januari 2007	Buana Asri Blok C No. 24
9	Novi Rahmawati	Ciamis, 4 Februari 2007	Sirojul Munir, RT02/RW07
10	Arifsya Ananka Niyegas	Bogor, 25 Januari 2007	Nanggawer Mekar, RT02/RW06
11	Faiz Fadhillah	Bogor, 16 Februari 2008	Bogor Asri Blok B No. 1
12	Khalif Fatur Syahdat	Jakarta, 24 Agustus 2007	Taman Nusa Citra
13	Kirana Natania Irawati	Bogor, 25 Mei 2008	Taman Bukit Asri Blok B2 No. 8

Tabel 2 Data Telah Diolah

No.	Nama	Tempat, Tanggal Lahir	Kelas	
			A	B
1	Dwi Novia Syafitri	Depok, 3 November 2006		√
2	Muhammad Adrian Eka S.	Cibinong, 9 Agustus 2007	√	
3	Kaniya Rahmah Ilahi	Bogor, 1 November 2007	√	
4	Maria Aulia Sianturi	Depok, 2 Agustus 2008	√	
5	Moreno Septianto	Bogor, 3 September 2007	√	
6	Dedy Maulana Iskandar	Ngawi, 23 Januari 2007	√	
7	Sri Wulan Dari	Bogor, 10 September 2006		√
8	Oliv Putri Vianti	Banyumas, 9 Januari 2007	√	
9	Novi Rahmawati	Ciamis, 4 Februari 2007	√	
10	Arifsya Ananka Niyegas	Bogor, 25 Januari 2007	√	
11	Faiz Fadhillah	Bogor, 16 Februari 2008	√	
12	Khalif Fatur Syahdat	Jakarta, 24 Agustus 2007	√	
13	Kirana Natania Irawati	Bogor, 25 Mei 2008	√	

Hasil dan Pembahasan

Data yang diperoleh dari operator TK Sarinah Bogor merupakan data yang *valid*. Data mencakup jumlah peserta didik yang terdaftar dalam kelas A dan kelas B tahun ajaran 2012 – 2022.

Tabel 3 Gabungan Deskripsi Data TK Sarinah Bogor Tahun Ajaran 2012 – 2022

No.	Tahun Ajaran	Kelas	
		A	B
1	2012/2013	11	2
2	2013/2014	16	11
3	2014/2015	12	5
4	2015/2016	1	0
5	2016/2017	6	21
6	2017/2018	2	3
7	2018/2019	9	1
8	2019/2020	0	10
9	2020/2021	4	10
10	2021/2022	0	4
11	2022/2023	18	7
Total		79	74

Analisis Data

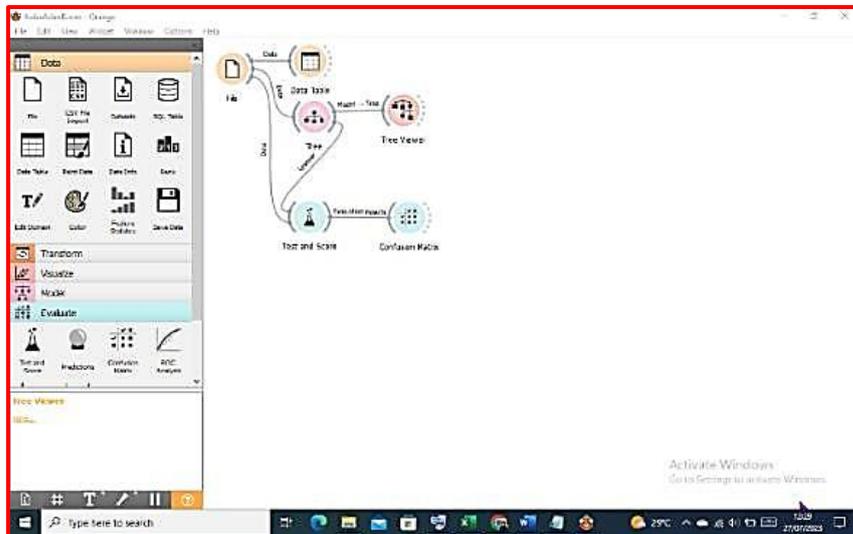
Tabel 3 merupakan tabel yang berisi informasi mengenai jumlah peserta didik kelas A dan kelas B untuk setiap tahun ajaran. Total keseluruhan pada Tabel 3 menunjukkan bahwa TK Sarinah Bogor memiliki total 79 peserta didik di kelas A dan 74 peserta didik di kelas B selama 11 tahun ajaran yang tercatat (2012/2013 hingga 2022/2023).

Uji Validitas

1. Orange

Firda Amalia dkk

Klasifikasi Peserta Didik
Menggunakan Algoritma
C4.5

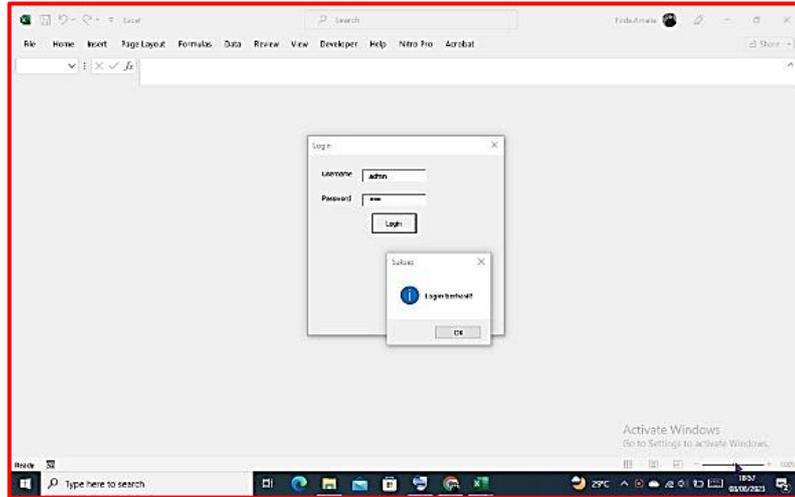


Gambar 1 Struktur Workflow

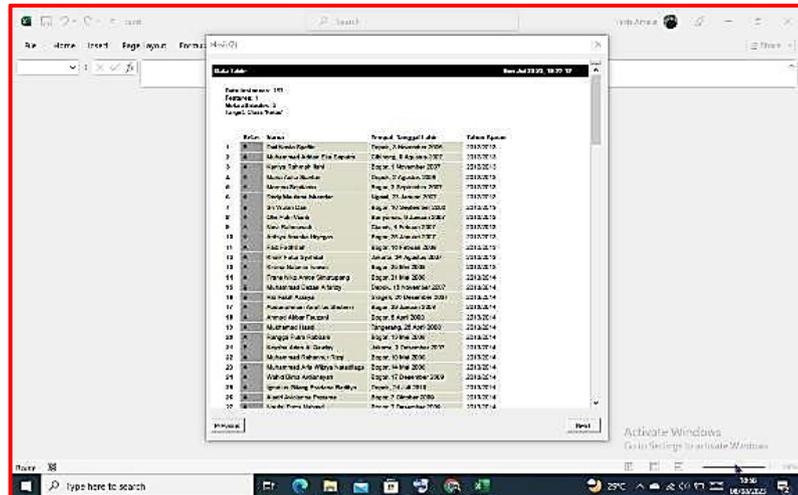
File	Type	Role	Values
1. No.	numeric	skip	
2. Nomor	categorical	target	A, B
3. Waktu	categorical	name	2013/01/13, 2013/01/14, 2014/2016, 20/ 5/2016, 2016/01/17, 2017/01/8, 2018/09/9, 20/ 8/2020, ...
4. Lokasi	text	name	
5. Tanggal, tanggal	text	name	

Gambar 2 Dataset TK Sarinah Bogor

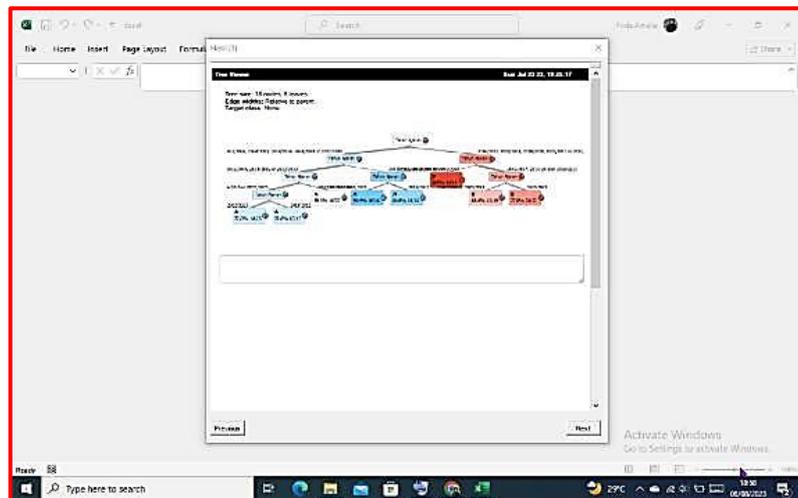
2. VBA



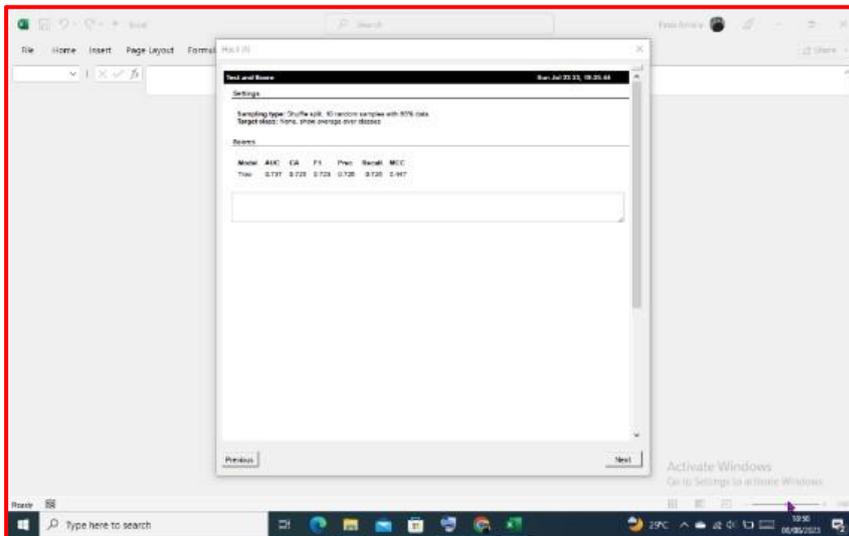
Gambar 3 Output Login



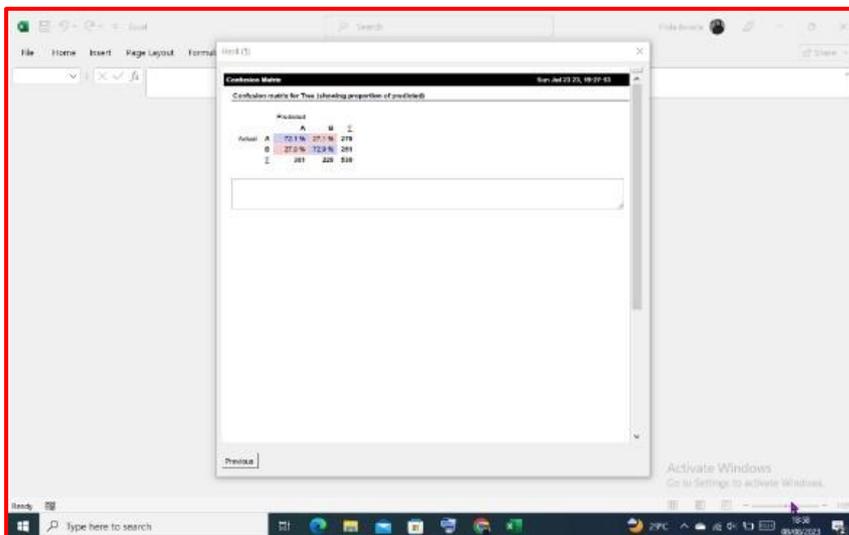
Gambar 4 Output Data Table



Gambar 5 Output Decision Tree



Gambar 6 Output Test & Score



Gambar 7 Output Confusion Matrix

Kesimpulan

Algoritma C4.5 adalah algoritma yang umumnya digunakan untuk membangun model klasifikasi berdasarkan pohon keputusan. Untuk mengonfirmasi apakah algoritma C4.5 cocok untuk membangun model klasifikasi peserta didik di TK Sarinah Bogor, perlu dilakukan penelitian lebih lanjut. Penting untuk mempertimbangkan data yang tersedia, ukuran sampel, serta kompleksitas masalah klasifikasi yang dihadapi di TK Sarinah Bogor sebelum memutuskan apakah algoritma ini sesuai atau tidak.

Hasil penelitian dengan menggunakan algoritma C4.5 dapat menjadi sumber informasi yang berharga dalam pengambilan keputusan terkait strategi pembelajaran di TK Sarinah Bogor. Algoritma ini dapat membantu mengidentifikasi faktor-faktor yang memengaruhi perkembangan individu peserta didik, sehingga dapat digunakan untuk

merancang strategi pembelajaran yang lebih efektif dan memberikan dukungan yang lebih sesuai bagi pengembangan individu peserta didik. Namun, keberhasilan implementasi hasil penelitian ini juga akan bergantung pada bagaimana hasil tersebut diintegrasikan ke dalam praktik pembelajaran di TK Sarinah Bogor.

Daftar Pustaka

- Anestiviya, V., & Pasaribu, A. F. (2021). Analisis Pola Menggunakan Metode C4.5 Untuk Peminatan Jurusan Siswa Berdasarkan Kurikulum (Studi Kasus : SMAN 1 Natar). *Jurnal Teknologi dan Sistem Informasi*, 2(1), 80-85.
- Azhari, D. Z., Damanik, I. S., & Suhendro, D. (2022). Penerapan Algoritma C4.5 Untuk Klasifikasi Tingkat Pemahaman Siswa Pada Mata Pelajaran Matematika. *FATIMAH: Penerapan Teknologi dan Sistem Komputer*, 1(1), 11-20.
- Irnanda, K. F., Hartama, D., & Windarto, A. P. (2021). Analisa Klasifikasi C4.5 Terhadap Faktor Penyebab Menurunnya Prestasi Belajar Mahasiswa Pada Masa Pandemi. *Jurnal Media Informatika Budidarma*, 5(1), 321-337.
- Jiang, J., Zhu, X., Han, G., & Gui, M. (2020). A dynamic trust evaluation and update mechanism based on C4. 5 decision tree in underwater wireless sensor networks. *IEEE Transactions on Vehicular Technology*, 69(8), 9031-9040.
- Nurindah, H. S., & Syahdan, S. (2022). Penerapan Algoritma C4.5 Dalam Klasifikasi Penerapan Peserta Didik Baru. *Jurnal Sains Benuanta*, 1(1), 1-6.
- Oktafiqurahman, A., Kusriani, & Nasiri, A. (2023). Implementasi Algoritma C4. 5 Untuk Tes Kepribadian Penerimaan Karyawan Di Dinas Perhubungan Provinsi Sulawesi Tengah. *Jurnal TIKomSiN (Teknologi Informasi dan Komunikasi)*, 11(1), 11-16.
- Purwanto, A., & Nugroho, H. W. (2023). Analisa Perbandingan Kinerja Algoritma C4. 5 Dan Algoritma K-Nearest Neighbors Untuk Klasifikasi Penerima Beasiswa. *Jurnal Teknoinfo*, 17(1), 236-243.

Firda Amalia dkk

Klasifikasi Peserta Didik
Menggunakan Algoritma
C4.5

HALAMAN INI SENGAJA DIKOSONGKAN