



IMPLEMENTASI MACHINE LEARNING DENGAN DECISION TREE ALGORITMA C4.5 DALAM PENERIMAAN KARYAWAN BARU PADA PT. GITAREKSA DINAMIKA JAKARTA

Imam Khoeri¹ dan Dadang Iskandar Mulyana²

STIKOM Cipta Karya Informatika Jakarta, Indonesia^{1 dan 2}
imamkhoeri2608@gmail.com¹ dan mahvin2012@gmail.com²

Diterima:

20 Juni 2021

Direvisi:

27 Juni 2021

Disetujui:

14 Juli 2021

Abstrak

Melakukan proses penerimaan karyawan merupakan hal yang sangat penting bagi suatu perusahaan, hal ini dilakukan mengingat karyawan dipandang sebagai salah satu aset penting bagi perusahaan dan perlu dikelola serta dikembangkan untuk mendukung suatu kelangsungan hidup, kemampuan untuk bersaing, mendapatkan laba, serta pencapaian pada tujuan perusahaan, PT. Gitareksa Dinamika Jakarta biasanya melakukan beberapa persyaratan atau kriteria untuk mengetahui kemampuan dan pribadi para pelamar tersebut, data hasilnya tersebut biasanya disimpan dalam suatu arsip yang harus dibandingkan satu persatu sehingga didapatkan suatu hasil atau keputusan. Tentu hal tersebut memakan waktu yang lama dan kurang efektif dalam mengambil suatu keputusan. Dengan demikian dibutuhkan suatu perencanaan model otomatis dari sekumpulan data, dengan tujuan memberikan komputer kemampuan untuk belajar (machine learning). Maka dari itu diperlukan sebuah metode untuk menentukan perancang dalam pengambilan keputusan penerimaan karyawan baru "Decision Tree Algoritma C4.5" tujuan khusus pada penelitian ini adalah membangun suatu sistem dengan mudah yang dapat mengelola data dan dapat diakses dengan mudah oleh staff HRD dalam menentukan karyawan baru pada PT. Gitareksa Dinamika Jakarta.

Kata kunci : *Machine learning; Decision tree; Algoritma C4.5; Penerimaan karyawan*

Abstract

Conducting the process of employee recruitment is very important for a company, this is done because employees are seen as one of the important assets for the company and need to be managed and developed to support a survival, ability to compete, earn profit, and achievement on the company's goals, PT. Gitareksa Dinamika Jakarta usually performs several requirements or criteria to know the ability and personal of the applicants, the data of the results are usually stored in an archive that must be compared one by one so that a result or decision is obtained. Of course it takes a long time and is less effective in making a decision. Thus it takes an automatic model planning of a data set, with the aim of giving the computer the ability to learn (machine learning). Therefore, a method is needed to determine the designer in the decision making of new employee acceptance "Decision Tree Algorithm C4.5" the specific purpose of this research is to build a system easily that can manage data and can be accessed easily by HRD staff in determining new employees at PT. Gitareksa Dynamics Jakarta.

Keywords: *Machine learning; Decision tree; C4.5 algorithm; Employee recipients*

PENDAHULUAN

Teknologi dan informasi berkembang dengan pesat sampai saat ini (Syamsuar & Reflianto, 2019) dan telah mampu membantu manusia dalam mengambil suatu keputusan (Limbong et al., 2020). Hal tersebut hampir menyentuh banyak bidang dan mempermudah kita dalam kegiatan sehari-hari yang kita lakukan (Nugroho & Al Fatta,

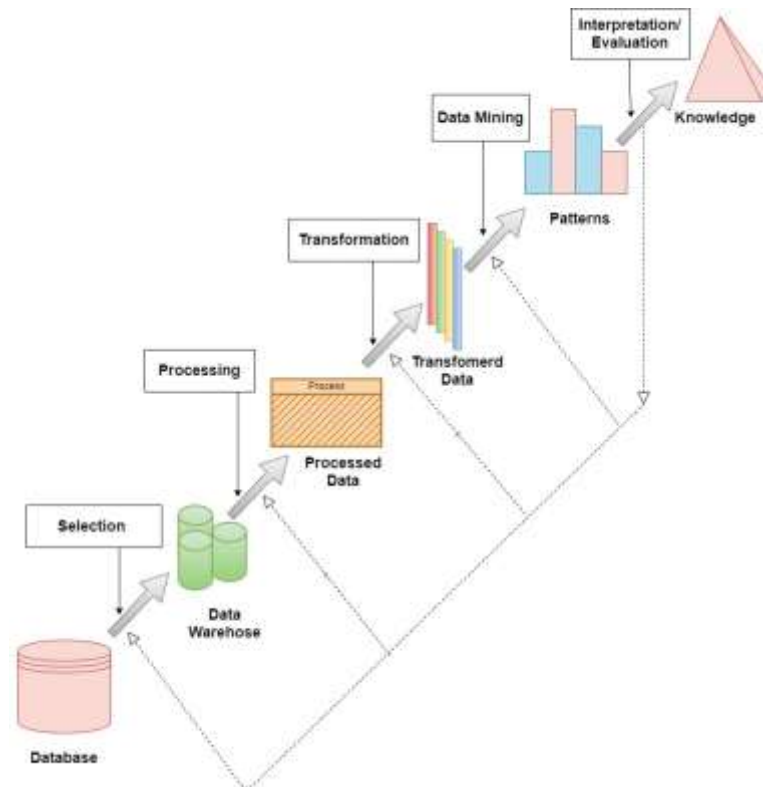
2015). Oleh karena itu, komputerisasi sistem merupakan salah satu sarana yang tepat untuk membantu pengambilan keputusan pada proses penerimaan karyawan baru (Sari, 2018). PT. Gitareksa Dinamika Jakarta merupakan salah satu perusahaan yang bergerak dalam pembiayaan dan penjualan mesin disel serta *service* pada mesin disel. Melakukan proses penerimaan karyawan merupakan hal yang sangat penting bagi suatu perusahaan (Sunandar & Satar, 2020), hal ini dilakukan mengingat karyawan dipandang sebagai salah satu aset penting bagi perusahaan (Amalia & Fakhri, 2016) dan perlu dikelola serta dikembangkan untuk mendukung suatu kelangsungan hidup (Lestari, 2017), kemampuan untuk bersaing, mendapatkan laba, serta pencapaian pada tujuan perusahaan (Suci, 2009).

Pada dasarnya, dalam suatu tujuan seleksi penerimaan karyawan adalah untuk mendapatkan sumber daya manusia yang tepat untuk suatu jabatan tertentu (Nuryanta, 2008), sehingga orang tersebut mampu bekerja secara optimal (Suwinardi, 2017) dan nantinya bekerja dan bertahan diperusahaan untuk jangka waktu yang lama (Lokollo & Syafruddin, 2013). Berdasarkan wawancara yang dilakukan bersama staff HRD perusahaan dalam melakukan penyeleksian karyawan baru PT. Gitareksa Dinamika Jakarta biasanya memberikan beberapa persyaratan atau kriteria untuk mengetahui kemampuan dan pribadi para pelamar tersebut, data hasilnya tersebut biasanya disimpan dalam suatu arsip yang harus dibandingkan satu persatu sehingga didapatkan suatu hasil atau keputusan. Tentu hal tersebut memakan waktu yang lama dan kurang efektif dalam mengambil suatu keputusan. Dengan demikian dibutuhkan suatu perencanaan model otomatis dari sekumpulan data, dengan tujuan memberikan komputer kemampuan untuk belajar yaitu *machine learning*. Maka dari itu diperlukan sebuah metode untuk menentukan perancangan pengambilan keputusan penerimaan karyawan baru yaitu "*Decision Tree Algoritma C4.5*".

Tujuan penelitian ini adalah membangun suatu sistem dengan mudah yang dapat mengelola data dan dapat diakses dengan mudah oleh staff HRD dalam menentukan karyawan baru pada PT. Gitareksa Dinamika Jakarta. Manfaat yang didapatkan tentunya untuk hasil yang diperoleh menjadi lebih efisien, akurat dan dapat mempersingkat waktu. Berdasarkan latar belakang yang telah dijabarkan di atas, maka di pilihlah untuk membangun suatu sistem cerdas dengan judul "*Implementasi Machine Learning dengan Decision Tree Algoritma C4.5 dalam Penerimaan Karyawan Baru pada PT. Gitareksa Dinamika Jakarta*" Diharapkan dengan penelitian ini, penulis dapat mampu memberikan solusi kepada masyarakat.

METODE PENELITIAN

Data mining adalah serangkaian proses yang memperkerjakan satu atau lebih teknik pembelajaran komputer untuk menganalisis dan mengekstrak pengetahuan secara otomatis atau serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Karena data mining adalah suatu rangkaian proses, data mining dapat dibagi menjadi beberapa tahap. Tahap-tahap tersebut bersifat interaktif di mana pemakai terlibat langsung atau dengan perantaraan *knowledge base* (Azwanti, 2018).



Gambar 1. Tahapan Data Mining

Pohon Keputusan (*Decision Tree*)

Pohon keputusan atau sering disebut *Decision Tree* adalah model *visual* untuk menyederhanakan proses pembuatan keputusan secara rasional. Visualisasi ini memungkinkan kita untuk memahami proses pembuatan keputusan yang terstruktur, bertahap, dan rasional. Pembuatan keputusan berarti memilih alternatif-alternatif keputusan yang tersedia. Merujuk pada pendapat Heizer dan Render (2005:326) yang dikutip oleh (Hadion Wijoyo, 2021). Konsep dari pohon keputusan adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan. Manfaat utama dari penggunaan pohon keputusan adalah kemampuannya untuk *break down* proses pengambilan keputusan yang kompleks menjadi lebih mudah sehingga pengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan (Hadion Wijoyo, 2021).

Algoritma C4.5

Metode penelitian yang digunakan adalah dengan *Decision Tree* Algoritma C4.5 dengan menggunakan algoritma pohon keputusan yang terkenal adalah C4.5. Pada akhir tahun 1970 sampai awal tahun 1980 J. Ross Quinlan, seorang peneliti di bidang *machine learning*, membuat sebuah algoritma *Decision Tree* yang dikenal dengan ID3 (*Iterative Dichotomiser*). Quinlan kemudian mengembangkan algoritma ID3 menjadi algoritma C4.5 yang merupakan penyempurnaan algoritma sebelumnya algoritma *Decision*. (Suntoro, 2019) Algoritma *Decision Tree* masuk ke dalam penerapan data mining klasifikasi (Kurniawan, 2018). Algoritma *Decision Tree* mengonstruksi pohon keputusan dari sebuah data training yang berupa *record-record* dalam basis data (Yuningsih et al., 2020). *Algoritma Decision Tree* banyak digunakan karena dapat eksplisit menggambarkan suatu pola atau pengetahuan atau informasi dalam bentuk pohon keputusan. *Algoritma Decision Tree* terdiri dari kumpulan *node* (simpul) yang dihubungkan oleh cabang, cabang tersebut bergerak ke bawah dari *root* (akar) *node* dan

berakhir di *leaf* (daun) *node*. *Leaf node* adalah *node* yang sudah tidak dapat dipecah lagi, *leaf node* merepresentasikan prediksi jawaban dari masalah (*data testing*). Pohon keputusan (*Decision Trees*) berbentuk terbalik, di mana *root node* berada di paling atas, sedangkan *leaf node* berada di paling bawah.

Konsep Entropi (S) yaitu merupakan jumlah bit yang diperkirakan dibutuhkan untuk dapat mengekstrak suatu kelas (+ atau -) dari sejumlah data acak pada ruang sampel S (Supriyanti et al., 2016). Entropi dapat dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas. Semakin kecil nilai *entropy* maka akan semakin *entropy* digunakan dalam mengekstrak suatu kelas (Riandari Fristi & Agustina Simangunsong, 2019). *Entropy* digunakan untuk mengukur ketidakaslisan S. Besarnya Entropi pada ruang sampel S didefinisikan dengan :

$$\text{Entropy}(S) = - \sum_{i=1}^n p_i \cdot \log_2 p_i$$

Dimana :

S : Himpunan kasus

A : Fitur

n : Jumlah partisi S

p_i : Proporsi dari S_i terhadap S

Konsep Gain (S,A) merupakan perolehan informasi dari atribut A relatif terhadap output data S. Perolehan informasi didapat dari output data atau variabel dependent S yang dikelompokkan berdasarkan atribut A, dinotasikan dengan gain (S,A) (Riandari Fristi & Agustina Simangunsong, 2019).

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \cdot \text{Entropy}(S_i)$$

Dengan :

S : Himpunan kasus

A : Atribut

n : Jumlah partisi atribut A

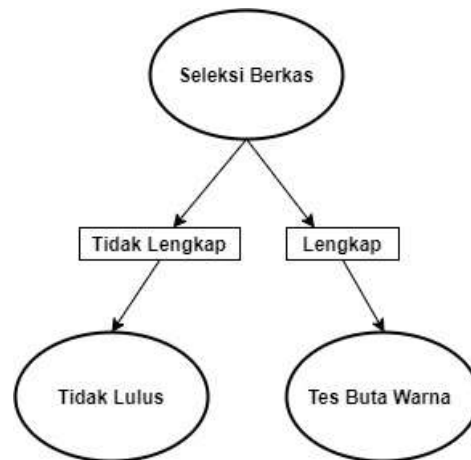
$|S_i|$: Jumlah kasus pada partisi ke i

$|S|$: Jumlah kasus dalam S

HASIL DAN PEMBAHASAN

Tabel 1. Perhitungan Node Setiap Pohon Keputusan

Atribut	Nilai	Jumlah Kasus	Lulus	Tidak Lulus	Entropy	Gain
Kelengkapan Berkas	Ya & Tidak	100	25	75	2	0.563117909
		83	25	58	1.731181183242	
Buta Warna	Ya & Tidak	17	0	17	0	0.336802784
		10	0	10	0	
Pengalaman Kerja	≥ 1 Tahun	90	25	65	1.847996907	0.265148445
	≤ 1 Tahun	30	15	15	1	
Tes Tertulis	≥ 70	70	10	60	2.807354922	0.358003696
	≤ 70	43	19	24	1.128337241	
		57	6	51	3.247927513	



Gambar 2. Pohon Keputusan Hasil Node 1

Tabel 2. Node 1

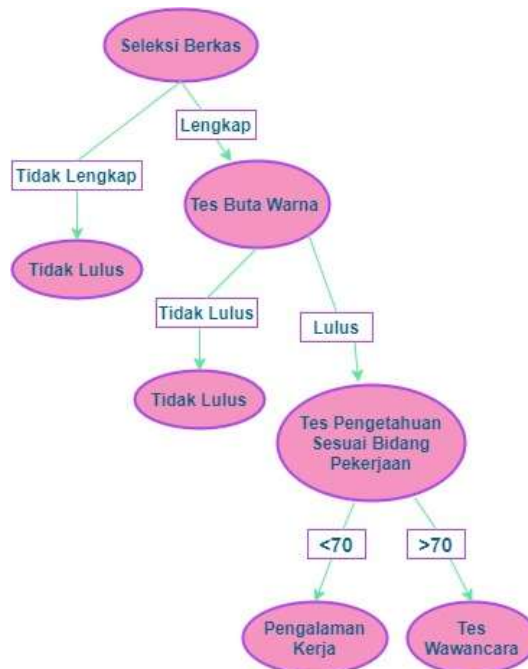
Atribut	Nilai	Jumlah Kasus	Lulus	Tidak Lulus	Entropy	Gain
Kelengkapan Berkas = Lengkap		83	25	58	1.731181183242	
Buta Warna	Ya	10	0	10	0	0.371476122
	Tidak	73	25	48	1.545968369	
Pengalaman Kerja	≥ 1 Tahun	27	15	12	0.847996907	0.22158566
	≤ 1 Tahun	56	10	46	2.485426827	
Tes Tertulis	≥ 70	43	19	24	1.178337241	0.198300194
	≤ 70	40	6	34	2.736965594	



Gambar 3. Pohon Keputusan Hasil Node 2

Tabel 3. Node 2

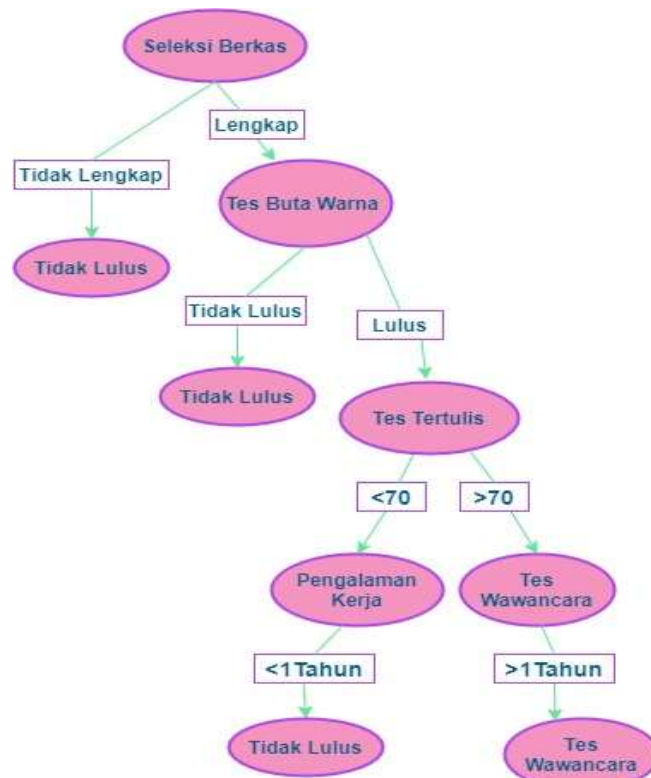
Atribut	Nilai	Jumlah Kasus	Lulus	Tidak Lulus	Entropy	Gain
Kelengkapan Berkas = Lengkap						
Buta Warna	Ya Tidak	73	25	48	1.545968369	
Pengalaman Kerja	≥ 1 Tahun	27	15	12	0.847996907	0.155005249
	≤ 1 Tahun	46	10	36	2.201633861	
Tes Tertulis	≥ 70	43	19	24	1.178337241	0.102337716
	≤ 70	30	6	34	2.321928095	



Gambar 3. Pohon Keputusan Hasil Node 2

Tabel 4. Node 3

Atribut	Nilai	Jumlah Kasus	Lulus	Tidak Lulus	Entropy	Gain
Kelengkapan Berkas = Lengkap						
Buta Warna = Tidak		30	6	24	2.3219281	
Tes Tertulis = ≤ 70						
Pengalaman Kerja	≥ 1 Tahun	9	6	3	0.847996907	2.1464393
	≤ 1 Tahun	21	0	21	0	



Gambar 3. Pohon Keputusan Hasil Node 3

Tabel 5. Node 4

No	Atribut	Kriteria
1	Seleksi Berkas	Lengkap
2	Test Buta Warna	Tidak
3	Pengalaman Kerja	$\leq 1\text{Tahun}/\geq 1\text{Tahun}$
4	Tes Pengetahuan Sesuai Bidang Pekerjaan	≥ 70
5	Tes Wawancara	≥ 70

No	Atribut	Kriteria
1	Seleksi Berkas	Tidak Lengkap
2	Test Buta Warna	Ya
3	Pengalaman Kerja	$\leq 1\text{Tahun}/\geq 1\text{Tahun}$
4	Tes Pengetahuan Sesuai Bidang Pekerjaan	≥ 70
5	Tes Wawancara	≥ 70

KESIMPULAN

Klasifikasi penerimaan calon karyawan PT. Gitareksa Jakarta menggunakan algoritma C4.5 yaitu dapat memberikan hasil keputusan sesuai dengan data yang sudah diinputkan dengan ketepatan dan keakuratan yang maksimal. Sistem ini menghasilkan keputusan seleksi secara objektif sesuai dengan data-data yang mengikuti tes seleksi. Mengurangi berkas dengan bentuk *hardcopy* atau menggunakan kertas karena bisa menyimpan di dalam *database* yang bisa sewaktu-waktu dibutuhkan bisa dibuka kembali. Sistem ini sudah bisa membantu dalam menentukan klasifikasi calon karyawan sesuai dengan kriteria yang telah ditentukan. Sistem ini dapat membantu SDM dalam mengelola data calon karyawan sehingga waktu yang dibutuhkan lebih efisien. Saran kepada PT. Gitareksa Jakarta dalam pengembangan aplikasi selanjutnya yaitu meningkatkan keamanan *web* agar lebih aman dari gangguan kejahatan internet dan membuat SMS *gateway* untuk memberi informasi kepada calon karyawan yang telah mendaftar.

BIBLIOGRAFI

- Amalia, S., & Fakhri, M. (2016). Pengaruh motivasi kerja terhadap kinerja karyawan pada PT. Gramedia Asri Media cabang Emerald Bintaro. *Jurnal Computech & Bisnis*, 10(2), 119–127.
- Azwanti, N. (2018). Analisa Algoritma C4.5 Untuk Memprediksi Penjualan Motor Pada Pt. Capella Dinamik Nusantara Cabang Muka Kuning. *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, 13(1)
- Hadion Wijoyo. (2021). *Teknik pengambilan keputusan*. Insan Cendekia Mandiri.
- Kurniawan, Y. I. (2018). Perbandingan Algoritma Naive Bayes dan C. 45 dalam Klasifikasi Data Mining. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIK)*, 5(4), 455–464.
- Lestari, M. (2017). Restrukturisasi Pendidikan Awal Perdamaian di Sekolah. *Prosiding Seminar Bimbingan Dan Konseling*, 1(1), 167–279.
- Limbong, T., Muttaqin, M., Iskandar, A., Windarto, A. P., Simarmata, J., Mesran, M., Sulaiman, O. K., Siregar, D., Nofriansyah, D., & Napitupulu, D. (2020). *Sistem Pendukung Keputusan: Metode & Implementasi*. Yayasan Kita Menulis.
- Lokollo, A., & Syafruddin, M. (2013). *Pengaruh Manajemen Modal Kerja Dan Rasio Keuangan Terhadap Profitabilitas Pada Industri Manufaktur Yang Terdaftar di Bursa Efek Indonesia (BEI) Tahun 2011*. Fakultas Ekonomika dan Bisnis, Undip.
- Nugroho, R. P. A., & Al Fatta, H. (2015). Teknologi Augmented Reality Sebagai Media Informasi Pemasaraan Pada Kupu-kupu Malam Car Auto-fashion. *Data Manajemen Dan Teknologi Informasi (DASI)*, 13(4), 29.
- Nuryanta, N. (2008). Pengelolaan Sumber Daya Manusia (Tinjauan Aspek Rekrutmen dan Seleksi). *EL TARBAWI*, 1(1), 55–69.
- Riandari Fristi & Agustina Simangunsong. (2019). *Penerapan Algoritma C4.5 Untuk Mengukur Tingkat Kepuasan Mahasiswa*. Cv. Rudang Mayang.
- Sari, F. (2018). *Metode dalam pengambilan keputusan*. Deepublish.
- Suci, R. P. (2009). Peningkatan kinerja melalui orientasi kewirausahaan, kemampuan manajemen, dan strategi bisnis (Studi pada Industri Kecil Menengah Bordir di Jawa Timur). *Jurnal Manajemen Dan Kewirausahaan*, 11(1), pp-46.
- Sunandar, H. S., & Satar, M. (2020). Tinjauan Tentang Proses Recruitment Tenaga Kerja di Kantor Pusat PT. Y Bandung. *Jurnal Industri Elektro Dan Penerbangan*, 6(1).
- Suntoro, J. (2019). *Data Mining Algoritma dan implementasi dengan pemrograman php*.

PT Elex Media Komputindo.

- Supriyanti, W., Kusriani, K., & Amborowati, A. (2016). Perbandingan Kinerja Algoritma C4. 5 dan Naïve Bayes Untuk Ketepatan Pemilihan Konsentrasi Mahasiswa. *Jurnal Informa: Jurnal Penelitian Dan Pengabdian Masyarakat*, 1(3), 61–67.
- Suwinardi, S. (2017). Profesionalisme dalam bekerja. *Orbith: Majalah Ilmiah Pengembangan Rekayasa Dan Sosial*, 13(2).
- Syamsuar, S., & Reflianto, R. (2019). Pendidikan dan tantangan pembelajaran berbasis teknologi informasi di era revolusi industri 4.0. *E-Tech: Jurnal Ilmiah Teknologi Pendidikan*, 6(2).
- Yuningsih, L., Setiawan, I. R., & Sunarto, A. A. (2020). Rancangan Aplikasi Prediksi Kelulusan Siswa Menggunakan Algoritma C4. 5. *Progresif: Jurnal Ilmiah Komputer*, 16(2), 121–132.



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.