JURNAL

IMPLEMENTASI ALGORITMA C4.5 DALAM PENENTUAN JURUSAN DI SMK PEMUDA PAPAR KEDIRI

C4.5 ALGORITHM IMPLEMENTATION IN DETERMINING THE DEPARTMENT OF SMK PEMUDA PAPAR KEDIRI

Oleh:

MARISA FITRI FATMAWATI

12.1.03.02.0273

Dibimbing oleh :

1. **RINI INDRIANI S.Kom., M.Kom**
2. **RISKY ASWI RAMADHANI M.Kom**

TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS NUSANTARA PGRI KEDIRI

2017

SURAT PERNYATAAN
ARTIKEL SKRIPSI TAHUN 2017

Yang bertanda tangan di bawah ini:

Nama Lengkap	: MARISA FITRI FATMAWATI
NPM	: 12.1.03.02.0273
Telepun/HP	: 087858919648
Alamat Surel (Email)	: marisafitrif@gmail.com
Judul Artikel	: Implementasi Algoritma C4.5 Dalam Penentuan Jurusan Di SMK Pemuda Papar Kediri
Fakultas – Program Studi	: Fakultas Teknik – Teknik Informatika
Nama Perguruan Tinggi	: UNIVERSITAS NUSANTARA PGRI KEDIRI
Alamat Perguruan Tinggi	: Jl. KH. Achmad Dahlan 76 Mojoroto, Kota Kediri

Dengan ini menyatakan bahwa :

a. artikel yang saya tulis merupakan karya saya pribadi (bersama tim penulis) dan bebas plagiarisme;
b. artikel telah diteliti dan disetujui untuk diterbitkan oleh Dosen Pembimbing I dan II.

Demikian surat pernyataan ini saya buat dengan sesungguhnya. Apabila di kemudian hari ditemukan ketidaksesuaian data dengan pernyataan ini dan atau ada tuntutan dari pihak lain, saya bersedia bertanggungjawab dan diproses sesuai dengan ketentuan yang berlaku.

Mengetahui		Kediri, 30 Januari 2017
Pembimbing I	Pembimbing II	Penulis,
Rini Indriani M.Kom. NIDN. 0725057003	Risky Aswi Ramadhani M.Kom NIDN. 0708049001	Marisa Fitri Fatmawati 12.1.03.02.0273

IMPLEMENTASI ALGORITMA C4.5 DALAM PENENTUAN JURUSAN DI SMK PEMUDA PAPAR KEDIRI

Marisa Fitri Fatmawati
12.1.03.02.0273
Fakultas Teknik - Teknik Informatika
marisafitrif@gmail.com
Rini Indriani S.Kom., M.Kom. dan Risky Aswi Ramadhani M.Kom.
UNIVERSITAS NUSANTARA PGRI KEDIRI

ABSTRAK

Setiap tahunnya cukup banyak siswa lulusan SMP maupun sederajat yang mendaftar di SMK Pemuda Papar Kediri. Namun banyak diantara mereka kurang matang dalam memilih jurusan yang ada sesuai dengan kemampuannya. Sehingga cukup banyak siswa baru yang merasa tidak cocok dengan jurusan yang telah dipilihnya, dan gagal di tengah jalan ketika mereka sudah diterima di SMK.

Perkembangan teknologi informasi saat ini dapat dimanfaatkan untuk melihat kemampuan siswa sehingga ketidakcocokan dan kebimbangan pemilihan jurusan dapat dikurangi dengan teknik pengelompokan data dari hasil data *mining*.

Penelitian ini menerapkan algoritma C4.5 ke dalam sistem penentuan jurusan, dimana dalam sistem ini data yang telah terkumpul akan diproses dalam tahapan proses *training* pengelompokan data, dan kemudian tahapan proses *testing*. Dengan menerapkan metode tersebut penelitian ini diharapakan dapat menghasilkan pola aturan penentuan jurusan, sehingga ketidakcocokan penentuan jurusan dapat diminimalisir.

Penelitian ini menghasilkan sistem Implementasi Algorima C4.5 Dalam Penentuan Jurusan Di SMK Pemuda Papar Kediri dengan dua *user* yang dapat mengoperasikan sistem, yaitu admin dan wakil kepala sekolah. Hasil yang diperoleh dari proses *testing* adalah keputusan jurusan yang sesuai dengan minat dan kemampuan siswa. Dengan adanya sistem Implementasi Penentuan Jurusan Di SMK Pemuda Papar Kediri, admin dapat melakukan proses penentuan jurusan dengan mudah dan cepat.

KATA KUNCI : data *mining*, algoritma C4.5, penentuan jurusan.

I. LATAR BELAKANG

SMK Pemuda Papar Kediri yang berlokasi di Jalan Pare No. 17 memiliki dua jurusan keahlian kompetensi yakni, jurusan APK (Administrasi Perkantoran), dan TKJ (Teknik Komputer dan Jaringan).

Setiap tahunnya cukup banyak siswa lulusan SMP maupun sederajat yang mendaftar di SMK Pemuda Papar Kediri. Namun banyak diantara mereka kurang matang dalam memilih jurusan yang ada sesuai dengan kemampuannya. Sehingga cukup banyak siswa baru yang merasa tidak cocok dengan jurusan yang telah dipilihnya, dan gagal di tengah jalan ketika mereka sudah diterima di SMK.

Pendekatan data mining dengan penerapan algoritma *Decision Tree* C4.5 akan dilakukan untuk menentukan jurusan yang akan diambil oleh siswa sesuai dengan

kemampuan, nilai akademik, nilai tes, dan minat siswa. Data yang telah terkumpul, selanjutnya akan dijadikan sebagai pembentuk pohon keputusan guna menemukan hubungan yang tersembunyi antara sejumlah calon *variable input* dan sebuah calon *variable* target.

Proses pada pohon keputusan adalah mengubah bentuk data berupa pohon keputusan menjadi model pohon, mengubah model pohon menjadi *rule*, dan menyederhanakan *rule* [1].

II. METODE

"*Data Mining* adalah analisis otomatis dari data yang berjumlah besar atau kompleks dengan tujuan untuk menemukan pola atau kecenderungan yang penting yang biasanya tidak disadari keberadaannya" [2].

Algoritma C4.5 bisa disebut juga algoritma pohon keputusan yang merupakan metode klasifikasi dan digunakan untuk memprediksi [1]. Langkah-langkah dalam membangun pohon keputusan:

1. pilih atribut sebagai akar
2. buat cabang untuk tiap-tiap nilai
3. bagi kasus dalam cabang
4. ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus seperti tertera dalam persamaan berikut [1]:

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

Keterangan :

S = himpunan kasus

A = atribut

n = jumlah partisi atribut A

$|S_i|$= jumlah kasus pada partisi ke-i

|S| = jumlah kasus dalam S

Perhitungan nilai *entropy* adalah perhitungan untuk mencari nilai yang nantinya dapat dijadikan cabang dengan nilai 0 (nol) atau nilai yang terendah dari semua atribut. Perhitungan *entropy* dapat dilihat pada persamaan berikut:

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

Keterangan :

S = himpunan kasus

n = jumlah partisi S

p_i = proporsi dari S_i terhadap S

Adapun atribut yang digunakan dalam penentuan jurusan adalah:

1. Nilai Ujian Nasional, dikelompok dengan variabel Baik (≥27.00), Cukup (≥21.00), Kurang (<21.00).
2. Nilai Komputer, dengan variabel Mampu (≥7.50), Kurang (<7.50).
3. Nilai Tes, dengan variabel Tinggi (≥8.00), Sedang (≥7.00), Rendah (<7.00).
4. Minat, dengan variabel APK, TKJ.

Data yang digunakan dalam penelitian ini adalah data siswa SMK Pemuda Papar tahun ajaran 2015/2016 dengan jumlah 200 data. Jumlah keseluruhan data dengan kriteria APK sebanyak 134 data, dan jumlah keseluruhan data dengan kriteria TKJ sebanyak 66 data. Berikut merupakan contoh perhitungan *entropy* pada node 1 penentuan jurusan:

$Entropy\ (TOTAL)$

$$= \left(-\frac{134}{200} * log_2\left(\frac{134}{200}\right)\right) + \left(-\frac{66}{200} * log_2\left(\frac{66}{200}\right)\right)$$

$= 0.91492$

Perhitungan nilai *Entropy* NUN :

$Entropy\ (NUN\ Baik)$

$$= \left(-\frac{95}{161} * log_2\left(\frac{95}{161}\right)\right) + \left(-\frac{66}{161} * log_2\left(\frac{66}{161}\right)\right)$$

$= 0.97647$

$Entropy\ (NUN\ Cukup)$

$$= \left(-\frac{35}{35} * log_2\left(\frac{35}{35}\right)\right) + \left(-\frac{0}{35} * log_2\left(\frac{0}{35}\right)\right)$$

$= 0$

$Entropy\ (NUN\ Kurang)$

$$= \left(-\frac{4}{4} * log_2\left(\frac{4}{4}\right)\right) + \left(-\frac{0}{4} * log_2\left(\frac{0}{4}\right)\right)$$

$= 0$

Perhitungan nilai Gain NUN :

$Gain\ (NUN)$

$$= 0.91492 - \left(\left(\frac{161}{200} * 0.97647\right) + \left(\frac{35}{200} * 0\right) + \left(\frac{4}{200} * 0\right)\right)$$

$= 0.12887$

Perhitungan *Gain* Komputer :

$Gain\ (Komp)$

$$= 0.91492 - \left(\left(\frac{165}{200} * 0.96729\right) + \left(\frac{35}{200} * 0.18718\right)\right)$$

$= 0.08416$

Perhitungan nilai *Gain* Tes:

$Gain\ (Tes)$

$$= 0.91492 - \left(\left(\frac{95}{200} * 0.9960\right) + \left(\frac{99}{200} * 0.7642\right) + \left(\frac{6}{200} * 0\right)\right)$$

$= 0.06351$

Perhitungan nilai *Gain* Minat :

$Gain\ (Minat)$

$$= 0.91492 - \left(\left(\frac{102}{200} * 0.28222\right) + \left(\frac{98}{200} * 0.95629\right)\right)$$

$= 0.30242$

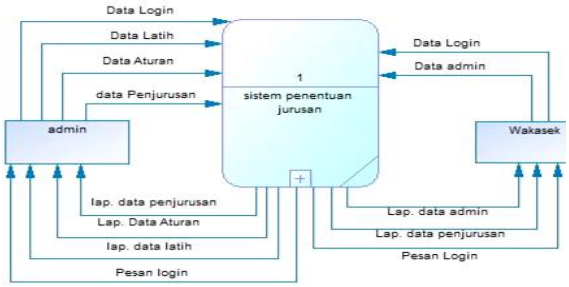
Diketahui dari contoh perhitungan node 1 di atas, nilai *gain* tertinggi adalah atribut Minat dengan nilai 0.30242, sehingga atribut Minat dapat dijadikan akar dari pohon keputusan.

III. HASIL DAN KESIMPULAN

A. Hasil Implementasi Sistem

1. Desain Proses DFD

Berikut merupakan Konteks Diagram Penentuan Jurusan SMK Pemuda Papar:



Gambar 1. Konteks Diagram

Sistem Penentuan Jurusan terdiri dari 2 *user* yakni, admin dan wakil kepala sekolah (wakasek). Kedua *user* dapat melakukan proses *login* dengan *username* dan *password* yang berbeda. Admin dapat melakukan proses *input* data latih, data aturan, dan data penjurusan. Sedangkan wakasek dapat meng*input*kan data admin, dan memperoleh data penjurusan.

2. *Input* Data Penjurusan

Gambar 2. Tampilan *Input* Data Penjurusan

Pada halaman ini, admin dapat menambah data siswa baru dan menyimpannya ke basisdata.

3. Tampilan Hasil Perhitungan Algoritma C4.5

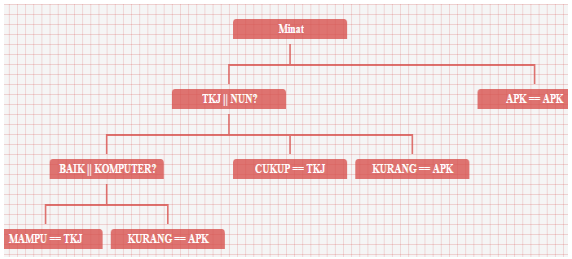
Node 1

Atribut	Nilai	Jumlah Kasus	APK	TKJ	Entropy	Gain
Total		200	134	66	0.91493	
nun	Baik	161	95	66	0.97647	0.12887
	Cukup	35	35	0	0	
	Kurang	4	4	0	0	
komputer	Mampu	165	100	65	0.96729	0.08416
	Kurang	35	34	1	0.18718	
tes	Sedang	99	77	22	0.7642	0.06351
	Tinggi	95	51	44	0.99608	
	Rendah	6	6	0	0	
minat	APK	102	97	5	0.28222	0.30242
	TKJ	98	37	61	0.95629	

Gambar 3. Tampilan Hasil C4.5

Pada halaman Algoritma C4.5, menampilkan hasil otomatis dari perhitungan algoritma C4.5. Data yang digunakan untuk perhitungan diambil dari data siswa. Pada halaman ini, admin tidak dapat menambah, *edit*, maupun hapus.

4. Tampilan Pohon Keputusan Dan *Rule*



Gambar 4. Pohon Keputusan C4.5

Pada halaman ini, menampilkan pohon keputusan dari hasil perhitungan C4.5.

Aturan Pohon Keputusan

No	Aturan
1	Jika minat tkj dan nun baik dan komputer mampu maka jurusan tkj
2	Jika minat tkj dan nun baik dan komputer kurang maka jurusan apk
3	Jika minat tkj dan nun cukup maka jurusan tkj
4	Jika minat tkj dan nun kurang maka jurusan apk
5	Jika minat apk maka jurusan apk

Gambar 5. *Rule* Pohon Keputusan

Admin dapat meng*edit* isi dari *rule*, sesuai dengan pohon keputusan yang ada.

5. Tampilan Data Penjurusan

Nama Siswa	Asal Sekolah	Nun	Komputer	Tes	Minat	Keputusan	C4.5
Sulistyorini	SMPN 2 PLEMAHAN	Baik	Mampu	Sedang	APK	TKJ	APK
Aziz Ramadani	SMPN 1 KAYEN KIDUL	Baik	Mampu	Sedang	APK	TKJ	APK
Septika Puspita Sari	SMPN 1 PAPAR	Baik	Mampu	Tinggi	TKJ	TKJ	TKJ
Sulaiman	MTsN MIFTAQUL HUDA	Baik	Kurang	Rendah	APK	APK	APK
Aditya Ruli Amanda	SMPN 2 PAPAR	Baik	Kurang	Sedang	TKJ	APK	APK
Pipit Wulandari	SMPN 2 PLEMAHAN	Baik	Kurang	Tinggi	APK	APK	APK
Meiliya Cahaya	SMPN 2 PAPAR	Cukup	Mampu	Sedang	APK	APK	APK
Rizka Febri Arinda	SMPN 1 PAPAR	Cukup	Mampu	Tinggi	TKJ	APK	APK
Rizza Sudi Humiyani	SMPN 2 PAPAR	Cukup	Kurang	Sedang	APK	APK	APK
Sukar Mila	SMPN 1 PAPAR	Kurang	Kurang	Rendah	TKJ	APK	APK

Gambar 6. Tampilan Data Penjurusan

Pada halaman ini, menampilkan data penjurusan siswa baru dari hasil *testing* dengan algoritma C4.5. Admin dapat menambah, meng*edit*, dan menghapus data penjurusan.

B. Kesimpulan

Berikut merupakan kesimpulan yang diperoleh dari hasil penelitian:

1. Telah dihasilkan sistem Implementasi Algorima C4.5 Dalam Penentuan Jurusan Di SMK Pemuda Papar Kediri dengan dua *user* yang dapat mengoperasikan sistem, yaitu admin dan wakil kepala sekolah. Hasil yang diperoleh dari proses *testing* adalah

keputusan jurusan yang sesuai dengan minat dan kemampuan siswa.

2. Dengan adanya sistem Implementasi Penentuan Jurusan Di SMK Pemuda Papar Kediri, admin dapat melakukan proses penentuan jurusan dengan mudah dan cepat.

IV. DAFTAR PUSTAKA

[1] Kusrini. Emha Taufiq Luthfi. 2009. *Algoritma Data Mining*. Yogyakarta: CV. ANDI OFFSET.

[2] Pramudiono, I. 2006. *Apa Itu Data Mining?*. Dalam Kusrini & Emha Tamrin Luthfi (Eds.), *Algoritma Data Mining* (hlm.3), Yogyakarta: CV. ANDI