

Penerapan Algoritma C4.5 Pada Aplikasi Bimbingan Karier Di SMK Telkom Purwokerto

Andi Dwi Riyanto¹⁾, Luzi Dwi Oktaviana²⁾, Umri Mardiyanti³⁾, Risa Christanti Gunardi⁴⁾
STMIK Amikom Purwokerto

Jl. Letjen Pol Soemarto (Watumas) Purwanegara Purwokerto Utara, 0281 623 321 / 0281 – 623 196
e-mail: andi@amikompurwokerto.ac.id¹⁾, oktaviana@amikompurwokerto.ac.id²⁾,
umti@amikompurwokerto.ac.id³⁾, risacg95@gmail.com⁴⁾

Abstrak

Bimbingan karier merupakan salah satu jenis layanan dari program bimbingan dan konseling (BK) yang ditujukan untuk membantu atau memfasilitasi peserta didik agar mencapai perkembangan yang optimal. Di SMK Telkom Purwokerto bimbingan karier membutuhkan analisis khusus dalam menentukan rekomendasi karier siswa setelah lulus. Metode yang digunakan dalam penelitian ini adalah teknik data mining klasifikasi dengan algoritme C4.5 disertai diskritisasi atribut. Dataset yang digunakan adalah data pribadi dan nilai Ujian Nasional alumni SMK Telkom Purwokerto. Hasil dari penelitian ini adalah didapatkan pohon keputusan dengan atribut paling berpengaruh adalah Nilai Ujian Nasional Bahasa Indonesia, hasil akurasi yang di dapat adalah 70,3414% dan aplikasi hasil penerapan pohon keputusan

Kata Kunci : *Bimbingan Karier, Algoritme C4.5, Diskritisasi, SMK Telkom Purwokerto.*

1. Pendahuluan

Dalam bidang pendidikan, bimbingan karier merupakan salah satu jenis layanan dari program bimbingan dan konseling (BK). Secara kelembagaan, bimbingan dan konseling adalah bagian dari keseluruhan program pendidikan di sekolah, yang ditujukan untuk membantu atau memfasilitasi peserta didik (siswa) agar mencapai perkembangan diri yang optimal [1]. SMK Telkom Purwokerto memberikan bimbingan karier bagi kelas XII yang akan segera lulus. Bimbingan karier ini dilakukan oleh guru BK. Menurut Bapak Lutfi selaku guru BK kelas XII SMK Telkom Purwokerto tahun ajaran 2015/2016, bimbingan karier dilakukan di kelas dan secara personal antara siswa dan guru BK. Guru BK akan memberi rekomendasi bagi siswa untuk melanjutkan kariernya setelah lulus dari SMK Telkom Purwokerto. Menurut Bapak Lutfi, selama ini bimbingan karier dilakukan tanpa ada analisis yang khusus, guru BK hanya menyebarkan angket yang berisi nama siswa dan minat untuk bekerja atau melanjutkan ke perguruan tinggi. Selama ini guru BK dalam memberi saran kepada peserta didiknya hanya berdasarkan opini yang di kiranya tepat untuk peserta didik tanpa ada analisis khusus.

SMK Telkom Purwokerto memiliki data alumni yang dapat dimanfaatkan untuk analisis penentuan rekomendasi bimbingan karier bagi kelas XII. Data yang sudah ada dapat digali dan diolah agar menjadi rekomendasi yang lebih berdasar jika dibandingkan dengan rekomendasi yang hanya perkiraan guru BK. Data alumni tersebut berupa biodata, nilai Ujian Nasional, jurusan semasa sekolah di SMK Telkom Purwokerto dan kegiatan di masa sekarang. Hasil dari 381 data alumni SMK Telkom Purwokerto menyimpulkan bahwa 52,49 % alumni melanjutkan bekerja, 1,83% alumni berwirausaha, 45,66% alumni melanjutkan ke perguruan tinggi. Dari data arsip tersebut dapat dilakukan penggalian data atau yang biasa disebut data mining. Hasil yang diperoleh dari proses data mining ini akan diterapkan dalam sebuah aplikasi. Harapannya setelah data alumni SMK Telkom Purwokerto diolah dengan data mining, bimbingan karier dapat dilakukan dengan akurat. Sehingga dapat menjadi pedoman bagi siswa yang kesulitan dalam menentukan pendidikan atau pekerjaannya setelah lulus dan memudahkan sekolah dalam memberi bimbingan karier bagi siswanya.

Menurut [2] istilah data mining digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Menurut[3], diskritisasi atribut bertujuan untuk mempermudah pengelompokan nilai berdasarkan kriteria yang telah ditetapkan. Hal ini juga bertujuan untuk menyederhanakan permasalahan dan meningkatkan akurasi. Menurut[4], diskritisasi dengan teknik binning adalah teknik pemisahan top-down berdasarkan pada jumlah tertentu bin. Metode ini juga

digunakan sebagai metode diskritisasi untuk reduksi data. Misalnya, nilai atribut dapat didiskritisasi dengan menerapkan sama-lebar atau sama-frekuensi, dan kemudian mengganti setiap nilai bindengan nilai median, seperti di *smoothingoleh median bin*. Menurut [1] bimbingan karier adalah suatu proses bantuan, layanan, pendekatan terhadap individu agar dapat mengenal dan memahami dirinya, mengenal dunia kerja, merencanakan masa depan yang sesuai dengan bentuk kehidupan yang diharapkannya, mampu menentukan dan mengambil keputusan secara tepat dan bertanggung jawab atas keputusan yang diambilnya itu sehingga mampu mewujudkan dirinya secara bermakna.

2. Metode Penelitian

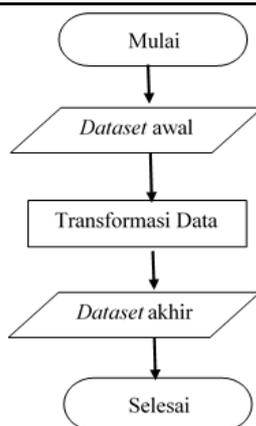
Menurut [2] Algoritme C4.5 merupakan algoritme yang digunakan untuk membentuk pohon keputusan. Pohon keputusan merupakan klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang mempresentasikan aturan - aturan dapat dengan mudah dipahami dengan bahasa alami. Dan mereka juga dapat diekspresikan dalam bentuk bahasa basis data seperti Structured Query Language untuk mencari record pada kategori tertentu. Pohon keputusan juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target. Data dalam pohon keputusan biasanya dinyatakan dalam bentuk tabel dengan atribut dan record. Atribut menyatakan suatu parameter yang dibuat sebagai kriteria dalam pembentukan pohon. Misalkan untuk menentukan main tenis, kriteria yang diperhatikan adalah cuaca, angin, dan temperatur. Salah satu atribut merupakan atribut yang menyatakan data solusi per item data yang disebut target atribut. Atribut memiliki nilai - nilai yang dinamakan dengan instance. Misalkan atribut cuaca memiliki instance berupa cerah, berawan, dan hujan. Proses pada pohon keputusan adalah mengubah bentuk data (tabel) menjadi model pohon, mengubah model pohon menjadi rule, dan menyederhanakan rule.

Pengumpulan data dimaksudkan untuk mendapatkan keterangan mengenai semua hal yang berhubungan dengan perancangan dan penerapan Sistem Pendukung Keputusan bimbingan karier menggunakan algoritme C4.5 di SMK Telkom Purwokerto. Dalam memperoleh data, digunakan beberapa metode, yaitu : Studi Pustaka, Wawancara dan Dokumentasi. Dalam penelitian ini digunakan 2 dataset yang terdiri dari 1013 objek data dan 864 objek data dimana data tersebut berisi 6 atribut kondisi dan 1 atribut keputusan. Atribut keputusan terdiri dari tiga nilai yaitu alumni bekerja, kuliah dan berwirausaha. Data tersebut diperoleh dari data siswa SMK Telkom Purwokerto. Atribut tersebut merupakan kriteria yang digunakan untuk menentukan kelanjutan siswa setelah lulus dari SMK Telkom Purwokerto.

Tabel 1 Atribut / Fitur yang digunakan pada dataset

No	Atribut	Keterangan	Nilai Atribut
1	UN Bahasa Indonesia	Berisikan nilai UN Bahasa Indonesia siswa	1-10
2	UN Bahasa Inggris	Berisikan nilai UN Bahasa Inggris siswa	1-10
3	UN Bahasa Matematika	Berisikan nilai UN Matematika siswa	1-10
4	UN Bahasa Kompetensi	Berisikan nilai UN Kompetensi siswa	1-10
5	Jenis Kelamin	Menjelaskan jenis kelamin siswa	Laki-laki, Perempuan
6	Jurusan di SMK	Menjelaskan jurusan siswa saat di SMK	RPL, TKJ, TJA
7	Kegiatan Sekarang	Menjelaskan kegiatan alumni saat ini apakah bekerja, kuliah atau berwirausaha	Bekerja, Kuliah, Berwirausaha

Preprocessing data merupakan tahapan yang dilakukan sebelum proses pengujian klasifikasi. Tahapan *preprocessing* data akan dijelaskan melalui Gambar 1.



Gambar 1 Tahap *Preprocessing* Data

Pada tahap ini dilakukan proses pengolahan data dengan tujuan mendapatkan data yang siap digunakan untuk penelitian. Yang pertama dilakukan dalam tahapan ini yaitu mendiskritisasi atribut dataset. Perubahan tersebut dilakukan untuk mempermudah pemahaman terhadap isi data.

2.1. Pembentukan pohon keputusan

Dalam tahap ini, dilakukan proses perhitungan nilai gain untuk sebuah instance dengan cara perhitungan nilai entropy lalu dihasilkan nilai gain tertinggi untuk di jadikan akar pohon. Setelah ditemukan akar pertama, maka proses lanjut ke penentuan akar kedua sampai kepada instance habis dan tidak ada akar lagi. Terbentuklah pohon keputusan berdasarkan nilai gain tertinggi.

2.2. Penerapan

Setelah dilakukan pengujian klasifikasi maka dilakukan penerapan algoritme kedalam sebuah aplikasi untuk menentukan kegiatan yang akan dilakukan_siswa_setelah_lulus.

2.3. Rekomendasi Karier

Melalui aplikasi yang telah dibangun, dapat terlihat hasil rekomendasi bimbingan karier dengan cara memasukan nilai atribut siswa yang membutuhkan rekomendasi.

3. Hasil dan Pembahasan

Dalam penelitian ini, data yang di dapat diolah terlebih dahulu untuk dapat diolah. Beberapa tahap pengolahan data adalah sebagai berikut :

3.1. Pre-processing Data

Dalam penelitian ini data yang digunakan diambil dari SMK Telkom Purwokerto berupa data pribadi siswa yang lulus pada tahun 2013-2016 yang diperoleh dari bagian Humas SMK Telkom Purwokerto, data nilai Ujian Nasional yang diperoleh dari bagian Kurikulum SMK Telkom Purwokerto. Pihak SMK Telkom Purwokerto merasa keberatan jika data yang bersifat privasi berupa nama dan di publikasikan. Penulis menggunakan samaran untuk data nama untuk menjaga privasi data siswa SMK Telkom Purwokerto.

Tabel 2 Data Pribadi Siswa yang Lulus pada Tahun 2013-2016

Nama Lengkap	Jenis Kelamin	Jurusan di SMK	Kegiatan sekarang
Siswa 1	P	RPL	kuliah
Siswa 2	L	RPL	kuliah
Siswa 3	P	RPL	kuliah
...
Siswa 380	P	TJA	bekerja
Siswa 381	P	TJA	kuliah

unbi = 4: krj (10.0)	unmat = 4: krj (0.0)
unbi = 5: krj (23.0/4.0)	unmat = 5: krj (0.0)
unbi = 6: krj (73.0/22.0)	unmat = 6: krj (0.0)
unbi = 7: kul (105.0/37.0)	unmat = 7: krj (1.0)
unbi = 8	unmat = 8: kul (2.0)
jur = TJA: krj (33.0/10.0)	unmat = 9: krj (0.0)
jur = TKJ: kul (19.0/1.0)	unmat = 10: krj (2.0)
jur = RPL: kul (27.0/4.0)	unkom = 8: krj (8.0/3.0)
unbi = 9: krj (48.0/15.0)	unkom = 9: krj (8.0/3.0)
unbi = 10	unkom = 10: kul (4.0/1.0)
unkom = 1: krj (0.0)	
unkom = 2: krj (0.0)	
unkom = 3: krj (0.0)	
unkom = 4: krj (0.0)	
unkom = 5: krj (3.0)	
unkom = 6	
uning = 1: kul (0.0)	
uning = 2: kul (0.0)	
uning = 3: kul (0.0)	
uning = 4: kul (0.0)	
uning = 5: kul (0.0)	
uning = 6: kul (0.0)	
uning = 7: krj (3.0/1.0)	
uning = 8: kul (3.0)	

Number of Leaves : 39
 Size of the tree : 44
 Algoritme pohon keputusan diatas
 memiliki nilai akurasi sebesar 70.3412%.

3.2. Penerapan

Tahap selanjutnya adalah penerapan kedalam sebuah aplikasi. Aplikasi yang dibangun berbasis dekstop dan berikut adalah flowchart dari aplikasi ini. Gambar 4.2 menjelaskan flowchart dari penerapan aplikasi.



Gambar 3 Flowchart dan Penerapan Aplikasi



Gambar 4 Tampilan Aplikasi

Pada tampilan awal aplikasi seperti pada gambar 4, pengguna dihadapkan pada kriteria penentu kegiatan selanjutnya setelah lulus. Hasil dari pengujian akan menampilkan output apakah siswa tersebut cocok melanjutkan kegiatan bekerja atau kuliah.

4. Simpulan

Berdasarkan hasil dari penelitian yang telah dilakukan, dapat disimpulkan hasil dari penelitian ini adalah sebagai berikut :

1. Didapatkan pohon keputusan dengan atribut paling berpengaruh adalah Nilai Ujian Bahasa Indonesia.
2. Dihasilkan aplikasi hasil penerapan dari pohon keputusan.

Agar penelitian ini dapat berkembang, berikut adalah saran yang diusulkan :

1. Memperbanyak jumlah *record dataset* dengan menghubungi alumni yang masih belum terjangkau.
2. Aplikasi dapat dikembangkan lebih jauh, di dalam aplikasi bisa ditambahkan database siswa sehingga lebih mudah digunakan. Aplikasi bisa ditambahkan fitur realisasi untuk mengetahui sejauh mana rekomendasi yang dihasilkan sesuai dengan keadaan nyata yang terjadi.

Daftar Pustaka

- [1] Supriatna, Mamat dan Nandang Budiman. 2012. Bimbingan Karir di SMK. From <http://file.upi.edu>, 13 Januari 2017
- [2] Kusrini dan Luthfi, E.T, 2009, Algoritma Data Mining Andi Offset Yogyakarta
- [3] Lesmana, I.P.D. 2012. Pengembangan Decision Tree J48 Untuk Diagnosis Penyakit Diabetes Mellitus. Politeknik Negeri Jember.
- [4] Han Jiawey, Kamber Micheline. 2012. Data Mining Concepts and Techniques. Elsevier Inc. USA.