

Algoritma Naïve Bayes Berbasis Backward Elimination untuk Prediksi Kesiapan Kerja Pada Siswa SMK

Husni Hidayat

Email: rajin_pangkal_pandai@hotmail.com

ITS NU Pekalongan

Abstrak

SMK (Sekolah menengah kejuruan) adalah kegiatan bimbingan belajar vokasi formal dalam mengembangkan bidang keilmuan, kemampuan kepribadian, sikap sosial, pembelajaran, pengembangan karir, perencanaan dan juga bakat siswa dalam memasuki dunia pekerjaan. Berdasarkan data yang ada pada SMK mutu lulusan SMK tidak semua siswanya memenuhi kualifikasi siap bekerja dikarenakan banyak faktor sehingga diperlukan sebuah pengolahan data yang menerapkan metode prediksi dengan teknik data mining. Salah satu teknik data mining adalah teknik klasifikasi yaitu *Naïve Bayes*, yang mampu menghasilkan nilai akurasi sampai 95.55% dalam memprediksi kesiapan kerja siswa SMK, Namun dari banyaknya faktor yang berpengaruh pada kesiapan kerja siswa SMK maka dibutuhkan penambahan fitur seleksi *Backward Elimination* yang mampu meningkatkan akurasi menjadi 96.95% dengan mengeliminasi beberapa fitur yang tidak relevan terhadap klasifikasi dan mendapat hasil yang lebih baik daripada menggunakan metode *Naïve Bayes* saja

Kata kunci : SMK; *naïve bayes*; *backward elimination*; prediksi

1. Pendahuluan

Pendidikan merupakan kegiatan mengoptimalkan perkembangan potensi, kecakapan, dan karakteristik pribadi peserta didik [1]. Salah satu bentuk pendidikan kejuruan adalah Sekolah Menengah Kejuruan. Jenjang pendidikan menengah ini memiliki tujuan utama mempersiapkan peserta didik terutama untuk bekerja dalam bidang tertentu [2]. Tumbuhnya minat bekerja tidak lepas dari pengaruh banyak faktor yang ada seperti lingkungan, keluarga, kondisi ekonomi [3]. Dunia usaha/industri sebagai pemakai lulusan dari dunia pendidikan juga memiliki peranan terhadap standar lulusan SMK, sehingga setelah lulus akan lebih mudah beradaptasi dengan dunia usaha/industri [1]. Kemitraan antara dunia usaha dan sekolah merupakan jalinan kerjasama yang dilakukan untuk memperoleh masukan atau keuntungan bagi kedua belah pihak yakni meningkatkan mutu bagi dunia kerja dan meningkatkan lulusan yang terserap dalam dunia kerja.

Dengan semakin majunya teknologi informasi pertumbuhan basis data yang berukuran besar dapat disimpan dan dikumpulkan tentu saja membutuhkan suatu teknik mengolah data yang menggunung

menjadi suatu informasi agar dapat memberikan manfaat guna mendukung pengambilan keputusan atau kebijaksanaan. Sekolah sebagai institusi pendidikan menyimpan data-data dalam jumlah yang besar, data-data tersebut dapat dianalisa menggunakan data mining untuk mengekstraksi dan menghasilkan pengetahuan yang memungkinkan bagi sekolah untuk memprediksi kesiapan kerja pada siswa SMK [4].

Dari penjelasan di atas, jelas bahwa melakukan prediksi kebutuhan dunia kerja dari SMK diperlukan sebuah metode prediksi dan juga merupakan hal yang penting bagi sekolah dimasa yang akan datang untuk menentukan kebijaksanaan strategis, dan sekolah bisa menggunakan mekanisme pendukung seperti pengawasan, nasihat, orientasi serta kebijakan strategis lainnya [5]. Maka pada penelitian ini untuk mengoptimalkan prediksi kesiapan kerja pada siswa SMK peneliti menggunakan *Naïve Bayes* dengan perhitungan probabilitas metode klasifikasi guna mengoptimalkan hasil prediksi dari 855 record data yang memiliki 21 atribut serta mengeliminasi atribut yang tidak signifikan maka dilakukan penggabungan *Naïve Bayes* berbasis

Backward Elimination. Metode *Backward Elimination* merupakan metode langkah mundur dalam membentuk sebuah model yang baik dengan meminimalkan/menghilangkan atribut yang secara signifikan tidak mempengaruhi model, kelebihan spesifik model penggabungan algoritma *Naive Bayes* dan *Backward Elimination* pada data set yang akan diteliti dibanding teknik-teknik diagnosis lain, *Backward Elimination* dapat mereduksi atribut data set sehingga dapat meningkatkan akurasi pada *Naive Bayes* [10].¹

2. Metode Penelitian

Penelitian ini menggunakan data sekunder dari jumlah 855 record, 543 record menghasilkan label Siap memasuki dunia kerja dan 312 record menghasilkan label Belum siap memasuki dunia kerja. Pada penelitian ini peneliti mencoba mengusulkan metode dengan tahap modeling untuk menyelesaikan prediksi kesiapan kerja pada siswa SMK menggunakan dua metode *Naive Bayes* dan *Backward Elimination*.

Setelah melakukan pengujian model menggunakan algoritma *Naive Bayes*, dimana dataset yang menggunakan *Backward Elimination* dan yang tidak menggunakan *Backward Elimination* akan dibandingkan akurasi dan efisiensinya menggunakan *Confusion matrix*. Peneliti melakukan proses ekperimental menggunakan aplikasi *RapidMiner 5.3.013*

Evaluasi dan validasi dilakukan dalam tahap ini menggunakan teknik yang ada dalam framework *rapidminer* versi 5.3.013, yaitu menggunakan teknik *Confusion Matrix* sebagai evaluasi untuk mengukur akurasi pada model yang digunakan, pengukuran dilakukan dengan menggunakan tabel klasifikasi yang bersifat prediktif sehingga didapatkan nilai *accuracy*, *precision* dan *recall*, untuk menganalisa kinerja dari algoritma dalam melakukan klasifikasi, serta *10-fold Cross Validation* untuk validasi model.

3. Hasil dan Pembahasan

Dalam eksperimen dan pengujian metode dilakukan uji coba terhadap metode *Naive Bayes* dan metode *Naive Bayes* berbasis

Backward Elimination, untuk mengetahui seberapa meningkat akurasi dari metode yang digunakan dengan melakukan evaluasi *confusion matrix*. Tahap selanjutnya adalah melakukan perbandingan model terbaik menggunakan algoritma *Naive Bayes* dan algoritma *Naive Bayes* berbasis *Backward Elimination*.

3.1. Pengujian *Naive Bayes*

Nilai probabilitas terdiri dari 2 Kelas/ label, yaitu: “Siap Bekerja” dan “Belum Siap Bekerja”, mempunyai 4 parameter: *Sangat Baik*, *Baik*, *Kurang*, *Cukup*, dari jumlah Atribut yang ada adalah 21 yaitu: *Motivasi Belajar*, *Hasil Pengalaman Praktik*, *Hasil Bimbingan Vokasi*, *Kondisi Ekonomi*, *Hasil Prestasi Belajar*, *Informasi Pekerjaan*, *Ekspektasi Masuk Kerja*, *Pengetahuan Umum*, *Tingkat Intelegensi*, *Bakat*, *Minat*, *Sikap*, *Nilai Nilai*, *Kepribadian*, *Keadaan Fisik*, *Penampilan Diri*, *Temperamen*, *Keterampilan*, *Kreativitas*, *Kemandirian* dan *Kedisiplinan*. Bisa dijadikan sebagai acuan dalam melakukan prediksi kesiapan kerja pada siswa SMK di masa mendatang. Berikut adalah sampel perhitungan nilai probabilitas dari atribut *Motivasi Belajar* :

Tabel I. Nilai Perhitungan Probabilitas
Motivasi Belajar

Parameter	Kelas			
	Siap Kerja	Belum Siap Kerja	Siap Kerja	Belum Siap Kerja
Sangat Baik	82	1		
Baik	424	89	543	312
Cukup	32	107		
Kurang	5	115		

Didapatkan nilai probabilitas dari atribut “*Motivasi Belajar*”, langkah selanjutnya membagi nilai dari setiap parameter dengan jumlah kelas yang sama. Total parameter “*Sangat Baik*” dengan Kelas “*Siap Kerja*” dibagi dengan jumlah kelas “*Belum Siap Kerja*”, berlaku juga untuk semua parameter dengan kelas yang sama dan semua atribut dalam dataset kesiapan kerja siswa SMK, maka mendapatkan hasil Sebagai berikut:

Tabel II. Nilai

Parameter	Siap Kerja	Belum Siap Kerja
Sangat Baik	82/543= 0,151	1/312= 0,003
Baik	424/543= 0,781	89/312= 0,285
Cukup	32/543= 0,059	107/312= 0,343
Kurang	5/543= 0,009	115/312= 0,369

3.2. Pengujian *Naïve Bayes* berbasis *Backward Elimination*

Pada tahap pengujian menggunakan *Backward Elimination* dari 21 atribut tereliminasi 5 atribut yang dianggap tidak relevan dan mempengaruhi tingkat akurasi. Hasil *Pengalaman Praktik, Bakat, Sikap, Keterampilan* dan *Kreativitas*. Dilihat dari nilai *weight* yang bernilai 0 artinya tidak berpengaruh sedangkan yang bernilai 1 berpengaruh terhadap hasil prediksi hal ini menjelaskan bahwa dalam perhitungan prediksi tingkat kesiapan kerja siswa SMK secara optimal dapat diprediksi melalui 16 atribut yang dipilih dan megeliminasi 5 atribut.

3.3. Pengujian Akurasi *Naïve Bayes*

Hasil pengujian *accuracy* dari eksperimen metode algoritma *Naïve Bayes* menggunakan validasi *10-Fold Cross validation* dengan *sampling type* menggunakan *linear sampling* diperoleh *accuracy* sebesar 95.55%, *precision* sebesar 96.02% dan *recall* sebesar 92.31% menggunakan tools *rapidminer 5.3.013*. seperti terlihat dalam tabel 4.6, tabel 4.7 dan Tabel III berikut:

Tabel III. Nilai Akurasi *Naïve Bayes*
Accuracy: 95.55%

	<i>true siap bekerja</i>	<i>true belum siap bekerja</i>	<i>class precision</i>
pred. siap bekerja	530	25	95.50%
pred. belum siap bekerja	13	287	95.67%
class recall	97.61%	91.99%	

$$\text{Nilai accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

$$= \frac{530+287}{530+287+13+25} = \frac{817}{855} \times 100\% = 95.55\%$$

3.2. Pengujian Akurasi *Naïve Bayes* berbasis *Backward Elimination*.

Hasil percobaan dengan algoritma *Naïve Bayes* berbasis *Backward Elimination*, menghasilkan akurasi sebesar 96.95% mengalami peningkatan sebesar 1.4% dibandingkan tanpa *Backward Elimination*. Juga menghasilkan 16 atribut yang terbaik yang bernilai satu dan 1 atribut dibuang yang bernilai nol.

Tabel IV. Nilai Akurasi *Naïve Bayes* berbasis *Backward Elimination*

Accuracy: 96.95%			
	<i>true siap bekerja</i>	<i>true belum siap bekerja</i>	<i>class precision</i>
pred. siap bekerja	532	15	97.26%
pred. belum siap bekerja	11	297	96.43%
Class recall	97.97%	95.19%	

$$\text{Nilai accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

$$= \frac{532+297}{532+297+11+15} = \frac{829}{855} \times 100\% = 96.95\%$$

4. Kesimpulan

Dalam penelitian ini dilakukan pengujian model dengan menggunakan *Naïve Bayes* dan *Naïve Bayes* berbasis *Backward Elimination* dengan menggunakan data kesiapan kerja siswa SMK sejumlah 855 record dengan 21 variabel. Model yang dihasilkan diuji untuk mendapatkan nilai *accuracy* dari setiap algoritma sehingga didapat pengujian dengan menggunakan data set dan setelah dilakukan pengujian didapat nilai *accuracy* adalah 95.55%. Sedangkan pengujian dengan *Naïve Bayes* berbasis *Backward Elimination* dilakukan eliminasi sebanyak 5 atribut yang tidak memberikan kontribusi, dari 21 variabel antara lain: “Hasil *Pengalaman Praktik*”, “*Bakat*”, “*Sikap*”, “*Keterampilan*” dan “*Kreativitas*”. Didapatkan hasil nilai *accuracy* 96.95%. Maka dapat disimpulkan pengujian data kesiapan kerja siswa SMK menggunakan *Naïve*

Bayes dan fitur seleksi *Backward Elimination* didapat metode tersebut lebih akurat dalam penentuan Prediksi kesiapan kerja siswa SMK dibandingkan dengan metode *Naïve Bayes* tunggal, ditandai dengan peningkatan nilai akurasi sebesar 1.40%.

5. Daftar Pustaka

- [1] D. Kurniasari dan G. Isnani (2015) “Analisis pelaksanaan kerjasama SMK dengan dunia usaha,” *Jurnal Pendidikan Bisnis dan Manajemen*, pp. 34-40, 2015.
- [2] A. Hidayati (2015), “Relevansi kompetensi lulusan sekolah menengah kejuruan dengan kebutuhan dunia usaha dan industri,” *Pendidikan Ekonomi & Bisnis*, pp. 1-12, 2015.
- [3] A. Fitriani, H. dan K (2012). , “Faktor-faktor yang mempengaruhi minat berwirausaha pada siswa kelas XII SMK Negeri 1 Kandeman Kabupaten Batang tahun 2011/2012,” *Jurusan Pendidikan Ekonomi*, pp. 1-5, 2012.
- [4] J. Han dan M. Kamber (2006), *Data Mining Concepts and Techniques*, San Fransisco: Diane Cerra, 2006
- [5] A. R. Rozzaqi (2015), “Naïve Bayes dan Filtering Feature Selection Information Gain untuk Prediksi Ketepatan Kelulusan Mahasiswa,” *Jurnal Informatika*, vol. I, pp. 30-41, 2015.
- [6] S. B. Martaleli dan S. Nugraha (2014), “Simulasi dan Prediksi Curah Hujan Mingguan Menggunakan Regresi Polinomial Dengan Metode Backward Elimination,” *Ilmu Komputer*, pp. 1-10, 2014..