ISSN: 2686-0260 September 2019 Hal:670-676

Penerapan Algoritma K-Nearest Neighbor untuk Prediksi Kelulusan Siswa pada SMK Anak Bangsa

Sekar Rizkya Rani¹, Sundari Retno Andani², Dedi Suhendro³
Program Studi Sistem Informasi, STIKOM Tunas Bangsa Pematangsiantar
Jln. Jendral Sudirman Blok A No. 1,2,3 Pematangsiantar
sekarrizkyarani@gmail.com

Abstract - Student graduation is very important for world education achievement, student graduation also influences the value of accreditation of an educational unit itself, therefor research on graduation prediction becomes a very interesting thing study, this study proposes the use of the K-Nearest Neighbor method to do predict student graduation at Anak Bangsa Private Vocational School. The result of the research is the value of k=5 with an accuracy rate of 93.55% which is determined as K-Optimal. The value of k=5 is applied to the K-NN algoritma to predict student graduation based on attendance, attitude, and value of knowledge.

Keywords: Datamining, Prediction of Vocational Student, K-Nearest Neigbor, SMK ANAK Bangsa

Abstrak - Kelulusan siswa merupakan hal sangat penting untuk prestasi dunia pendidikan, kelulusan juga berpengaruh terhadap nilai akreditasi suatu pendidikan itu sendiri, oleh karena itu penelitian tentang prediksi kelulusan menjadi hal yang sangat menarik untuk diteliti, Penelitian ini mengusulkan penggunaan metode K- Nearest Neighbor untuk melakukan prediksi kelulusan siswa pada SMK Swasta Anak Bangsa. Hasil penelitian yang diperoleh adalah nilai K=5 dengan tingkat akurasi sebesar 93,55 % yang ditetapkan sebagai K-Optimal. Nilai K=5 diterapkan pada algoritma K-Nearest Neighbor untuk prediksi kelulusan siswa berdasarkan kehadiran, sikap dan nilai pengetahuan.

Kata Kunci: Data Mining, Prediksi Kelulusan Siswa, K-Nearest Neighbor, SMK Anak Bangsa

1. PENDAHULUAN

Dalam meningkatkan Sumber Daya Manusia (SDM), bidang pendidikan memiliki peranan penting untuk kemajuan suatu bangsa. Salah satu cara untuk menilai keberhasilan lembaga pendidikan untuk menghasilkan individu-individu yang berkualitas dapat dilihat dari nilai ujian siswa yang bersangkutan. Ujian nasional merupakan salah satu alat evaluasi yang dikeluarkan pemerintah untuk mengukur tingkat pencapaian tujuan pendidikan yang telah ditetapkan [1] Dalam lembaga pendidikan, ujian nasional digunakan untuk menentukan kelulusan dalam tingkatan pendidikan dan menjadi tolak ukur satuan pendidikan. Bagi pihak sekolah terutama para guru, tingkat kelulusan siswa menjadi prioritas utama karena berkaitan dengan akreditasi sekolah. Untuk mampu bersaing dan memiliki kualitas sekolah yang baik, upaya yang dilakukan pihak sekolah adalah meminimalisir angka ketidaklulusan siswa dengan cara melakukan proses pengolahan terhadap data kelulusan menggunakan data mining. Data mining sangat diperlukan untuk mengolah jumlah data yang besar terutama data kelulusan siswa yang setiap tahun terus bertambah sehingga informasi yang

Algoritma K-Nearest Neighbor untuk Prediksi Kelulusan Siswa pada SMK (Sekar Rizkya Rani) | 670

ISSN: 2686-0260 September 2019 Hal:670-676

dihasilkan menjadi lebih akurat dalam perkembangan saat ini dan waktu mendatang.

Salah satu jurnal yang menjadi referensi dalam penulisan skripsi ini yaitu penelitian yang dilakukan oleh [2]. Dalam jurnal tersebut, peneliti melakukan penelitian dalam memprediksi predikat mahasiswa. Atribut yang digunakan dalam proses prediksi yaitu Jenis Kelamin, Jenis Tinggal, Umur, Jumlah Satuan Kredit Semester (SKS) dan Jumlah Nilai Mutu. Jurnal lain yang menjadi acuan adalah jurnal tentang Prediksi tingkat pengunduran diri mahasiswa mengunakan metode K-Nearest Neighbor yaitu penelitian yang dilakukan oleh [3] yang membahas tentang Prediksi Tingkat Pengunduran Diri Mahasiswa. Atribut yang digunakan yaitu IPK, Pekerjaan Orang Tua, Jurusan, dan Semester. Penelitian selanjutnya yang menjadi referensi yaitu penelitian yang dilakukan oleh [4] yang membahas tentang Prediksi Penjualan Barang. Adapun data yang digunakan adalah data di bulan sebelumnya. Berdasarkan latar belakang diatas, penulis mengangkat judul "Penerapan Algoritma K-Nearest Neighbor untuk Prediksi Kelulusan Siswa", diharapkan hasil penelitian ini dapat memberikan masukan kepada Pihak Sekolah khususnya SMK Swasta Anak Bangsa dalam mempertimbangkan kelulusan siswasiswi secara individu maupun keseluruhan serta dapat membantu pihak sekolah dalam mengurangi jumlah ketidak lulusan siswa agar kualitas kelulusan siswa menjadi lebih baik.

2. METODOLOGI PENELITIAN

Dalam melakukan penelitian ini, penulis melakukan analisa data statistik deskriptif. Pada tahap ini dilakukan pengumpulan data kelulusan siswa yang digunakan untuk penelitian ini adalah wawancara, observasi dan studi pustaka. Penelitian ini dilakukan dengan cara melakukan peninjauan dan pencatatan langsung ke sekolah SMK Swasta Anak Bangsa.

2.1. Prediksi Kelulusan

"Prediksi adalah suatu proses atau tahap untuk memperkirakan , ramalan ataupun perkiraan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan" (Wikipedia, 2018). Prediksi kelulusan adalah suatu perkiraan untuk melihat lulus tidaknya siswa dalam ujian nasional dengan standar kelulusan tertentu.

2.2. Algoritma K-Nearest Neighbor

K-Nearest Neighbor (KNN) merupakan metode melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Metode ini bertujuan untuk mengklasifikasikan objek baru berdasarkan atribut dan training sample. Diberikan suatu titik query, selanjutnya akan ditemukan sejumlah K objek atau titik training yang paling dekat dengan titik query. Nilai prediksi dari query akan ditentukan berdasarkan klasifikasi tetanggaan [5]. K-Nearest Neighbor memiliki keunggulan diantaranya "pelatihan sangat cepat, sederhana dan mudah dipelajari, efektif jika data pelatihan besar" [6]. Langkah-langkah untuk menghitung algoritma K-NN:

1. Menentukan nilai k.

- 2. Menghitung kuadrat jarak *euclid* (*query instance*) masing-masing objek terhadap *training* data yang diberikan. Persamaan (1) menunjukkan perhitungan kuadrat jarak *euclid* (*query instance*) masing-masing objek terhadap *training* data yang diberikan.
- 3. Kemudian mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak *euclid* terkecil
- 4. Menghitung rata-rata dari nilai objek pada jangkauan K dengan menggunakan kategori *Nearest Neighbor* yang terdekat (jangkauan K), maka dapat diprediksi nilai *query instance* yang telah dihitung.

3. HASIL DAN PEMBAHASAN

3.1. Analisa

Data penelitian diperoleh dari Sekolah SMK Swasta Anak Bangsa yaitu data kelulusan siswa tahun 2017/2018 dengan jumlah siswa 60 orang. Data yang digunakan terdiri dari 3 kriteria yaitu Nilai Pengetahuan, Sikap dan Kehadiran Berikut adalah data penelitian yang digunakan, dapat dilihat pada tabel 1:

Tabel 1. Data Training

No.	Nama Peserta	NRR	Sikap	Kehadiran	KET
110.	Numu i esercu	14144	этар	nendan di	ILL I
1	Al-Friyansyah Damanik	77.05	3.00	91.00	Lulus
2	Al-Hasanul Bisri Dmk	76.96	3.00	93.00	Lulus
3	Ambri Saragih	81.27	3.00	96.00	Lulus
4	Anggi Febriani	78.92	3.00	92.00	Lulus
5	Ayu Oktaviani	77.78	4.00	94.00	Lulus
6	Cisi Elvira	79.73	4.00	95.00	Lulus
7	Desi Fitriani	80.29	4.00	92.00	Lulus
8	Dikki Abbdul Kahar	77.30	3.00	93.00	Lulus
9	Diky Ardiansyah	80.64	3.00	94.00	Lulus
10	Elly Syafitri	81.47	4.00	92.00	Lulus
11	Fitriani	61.60	2.00	70.00	Tidak Lulus
12	Ika Sintya	80.11	3.00	92.00	Lulus
13	Ilham Kurniadi	77.59	4.00	91.00	Lulus
14	Kiki Sundari	61.27	2.00	69.00	Tidak Lulus
15	Manda Sari	83.51	4.00	97.00	Lulus
16	Mawan Ardiansyah	78.36	3.00	92.00	Lulus
17	Mayang Sari	82.51	4.00	95.00	Lulus
			•••	•••	•••
57	Wana Nopia	79.71	4.00	95.00	Lulus
58	Widiya Aryuning	80.44	4.00	95.00	Lulus
59	Widodo Cahyandi	79.28	4.00	93.00	Lulus
60	Yogi Kusuma	78.45	4.00	94.00	Lulus

Tabel 2. Data Testing

No.	Nama Peserta	NRR	Sikap	Kehadiran	KET
1	Adinda Fitriyani	80.64	3.00	93.00	Lulus
2	Andryansyah	81.47	4.00	91.00	Lulus
3	Anggiat Ramadhan	80.64	3.00	92.00	Lulus
4	Arthika Pratiwi	80.11	3.00	93.00	Lulus
5	Asni	77.59	4.00	92.00	Lulus

No.	Nama Peserta	NRR	Sikap	Kehadiran	KET
6	Azwar Zuelpa Purba	78.50	3.00	93.00	Lulus
7	Dimas Anggoro	83.51	4.00	93.00	Lulus
8	Dimas Ramadani	82.86	4.00	91.00	Lulus
9	Heri Suriadi Saragih	75.00	3.00	92.00	Lulus
10	Iqbal Pratama	81.36	3.00	92.00	Lulus
			:	***	
28	Sri Rahayu	80.00	3.00	94.00	Lulus
29	Supia Putri	80.00	4.00	97.00	Lulus
30	Susanti Saragih	82.86	4.00	94.00	Lulus
31	Syahidul	75.00	3.00	92.00	Lulus

3.2. Penyelesaian Algoritma algoritma K-Nearest Neighbor

Langkah – langkah dalam menyelesaikan perhitungan manual data mining menggunakan K-Nearest Neighbor yaitu:

- 1. Nilai k yang digunakan adalah 5.
- 2. Menghitung kuadrat jarak *Euclid* (*query instance*) masing-masing objek terhadap sampel data yang diberikan dengan menggunakan rumus :

di =
$$\sqrt{\sum_{i=1}^{p} (x2i - x1i)^2}$$

Untuk objek pertama pada data *training* dihitung jaraknya dengan data testing menggunakan *euclidien distance* sebagai berikut :

$$d = \sqrt{(77.05 - 80.64)^2 + (3.00 - 3.00)^2 + (91.00 - 93.00)^2} = \sqrt{16.8881}$$

$$= 4.1095$$

$$d = \sqrt{(76.96 - 80.64)^2 + (3.00 - 3.00)^2 + (93.00 - 93.00)^2} = \sqrt{13.5424}$$

$$= 3.68$$

Tabel 3. Hasil Perhitungan Ouery Distance

No.	Nama Peserta	NRR	Sikap	Kehadiran	Query Distance
1	Al-Friyansyah Damanik	77.05	3.00	91.00	4.109513353
2	Al-Hasanul Bisri Damanik	76.96	3.00	93.00	3.68
3	Ambri Saragih	81.27	3.00	96.00	3.065436347
4	Anggi Febriani	78.92	3.00	92.00	1.989572818
5	Ayu Oktaviani	77.78	4.00	94.00	3.190548542
6	Cisi Elvira	79.73	4.00	95.00	2.414145812
7	Desi Fitriani	80.29	4.00	92.00	1.456880228
8	Dikki Abbdul Kahar	77.30	3.00	93.00	3.34
9	Diky Ardiansyah	80.64	3.00	94.00	1
10	Elly Syafitri	81.47	4.00	92.00	1.639786571
59	Widodo Cahyandi	79.28	4.00	93.00	1.688075828
60	Yogi Kusuma	78.45	4.00	94.00	2.606933064

3. Kemudian mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak Euclid terkecil.

Tabel 4. Mengurutkan Objek ke Jarak Euclidien Terkecil

10.001 111 10118011 01011111 0 0)011 110)			J =			
No.	Nama Peserta	NRR	Sikap	Kehadiran	Jarak Terkecil	
1	Al-Friyansyah Damanik	77.05	3.00	91.00	43	
2	Al-Hasanul Bisri Damanik	76.96	3.00	93.00	38	
3	Ambri Saragih	81.27	3.00	96.00	26	
4	Anggi Febriani	78.92	3.00	92.00	13	
5	Ayu Oktaviani	77.78	4.00	94.00	30	
6	Cisi Elvira	79.73	4.00	95.00	16	
7	Desi Fitriani	80.29	4.00	92.00	9	
8	Dikki Abbdul Kahar	77.30	3.00	93.00	33	
9	Diky Ardiansyah	80.64	3.00	94.00	1	
10	Elly Syafitri	81.47	4.00	92.00	11	
59	Widodo Cahyandi	79.28	4.00	93.00	12	
60	Yogi Kusuma	78.45	4.00	94.00	20	

4. Mengumpulkan label class Y (klasifikasi *Nearest Neighborhood*).

Pada proses ke-4 ini, hasil dari pengelompokkan "apakah termasuk *Nearest Neighborhood* (k) akan dikelompokkan dengan "Y=Klasifikasi"

Tabel. 5. Y = Kategori Nearest Neighbor (k)

No.	Nama Peserta	NRR		Kehadiran	<u> </u>	Y=
NO.	Nama Peserta	NKK	Sikap	Kenaunan	Jarak	_
					Terkecil	Kategori
						Nearest
1	Al-Friyasyah Dmk	77.05	3.00	91.00	43	
2	Al-Hasanul Bisri	76.96	3.00	93.00	38	
	Damanik					
3	Ambri Saragih	81.27	3.00	96.00	26	
4	Anggi Febriani	78.92	3.00	92.00	13	
5	Ayu Oktaviani	77.78	4.00	94.00	30	
6	Cisi Elvira	79.73	4.00	95.00	16	
7	Desi Fitriani	80.29	4.00	92.00	9	
8	Dikki A Kahar	77.30	3.00	93.00	33	
9	Diky Ardiansyah	80.64	3.00	94.00	1	Lulus
10	Elly Syafitri	81.47	4.00	92.00	11	
				•••		
59	Widodo Cahyandi	79.28	4.00	93.00	12	
60	Yogi Kusuma	78.45	4.00	94.00	20	

Dapat dilihat pada tabel 5. mayoritas klasifikasi yang muncul adalah "LULUS", sehingga data testing yang ingin diklasifikasikan termasuk kedalam klasifikasi "Ya" atau lulus.

Tabel 6. Hasil Perhitungan Keseluruhan Data Testing

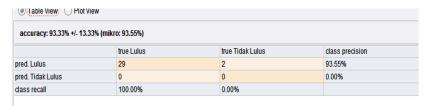
NO	Nama Peserta	NRR	Sikap	Kehadiran	KET
1	Adinda Fitriyani	80.64	3.00	93.00	Lulus
2	Andryansyah	81.47	4.00	91.00	Lulus

3	Anggiat Ramadhan	80.64	3.00	92.00	Lulus
4	Arthika Pratiwi	80.11	3.00	93.00	Lulus
5	Asni	77.59	4.00	92.00	Lulus
6	Azwar Z Purba	78.50	3.00	93.00	Lulus
7	Dimas Anggoro	83.51	4.00	93.00	Lulus
8	Dimas Ramadani	82.86	4.00	91.00	Lulus
9	Heri S Saragih	75.00	3.00	92.00	Lulus
10	Iqbal Pratama	81.36	3.00	92.00	Lulus
				•••	
29	Supia Putri	80.00	4.00	97.00	Lulus
30	Susanti Saragih	82.86	4.00	94.00	Lulus
31	Syahidul	75.00	3.00	92.00	Lulus

Dari tabel 6 dapat diperoleh hasil prediksi kelulusan siswa Menggunakan Algoritma *K-Nearest Neighbor*. 29 siswa diprediksi lulus dan 2 siswa diprediksi tidak lulus.

3.3. Hasil

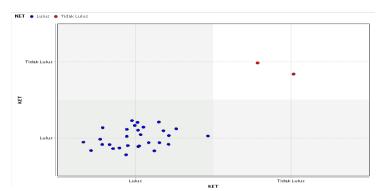
Hasil akhir data testing pada bagian pembahasan ini akan menjelaskan mengenai hasil akurasi yang didapatkan oleh data testing pada rapidminer 5.3. dapat dilihat pada gambar berikut ini



Gambar 1. Akurasi Data Testing

3.4. Pembahasan

Berdasarkan gambar 4.18 tersebut menunjukan bahwa akurasi dengan menggunakan algoritma *KNN* adalah sebesar 93,55% dengan kondisi k=5. Sehingga berdasarkan akurasi tersebut diketahui grafik hasil pada data testing dapat dilihat pada gambar berikut.



Gambar 2. Hasil Akhir Data Testing

Berdasarkan pada gambar 4.19. dapat diketahui bahwa pada klasifikasi lulus (merah) memiliki banyak node yaitu 29, sedangkan pada klasifikasi tidak lulus (biru) memiliki 2 node yaitu Novita Amelia dan Rindiani.

4. KESIMPULAN

Nilai K-Optimal pada algoritma KNN untuk prediksi kelulusan siswa adalah K=5. Dari proses K-fold Cross Validation didapatkan tingkat akurasi untuk K=5 pada algoritma KNN untuk prediksi kelulusan berdasarkan nilai pengetahuan, sikap, dan kehadiran adalah sebesar 93,55%. Dari 31 Siswa hasil prediksi kelulusan menggunakan algoritma K-Nearest Neighbor 29 siswa diprediksi lulus dan 2 siswa diprediksi tidak lulus.

DAFTAR PUSTAKA

- [1] U. F. Wiji setyaningsih, "DECISION SUPPORT SYSTEM PREDIKSI KELULUSAN SISWA SMK MENGGUNAKAN METODE NAIVE BAYES," vol. 2, 2015.
- [2] Mustakim and G. O. F, "Algoritma K-Nearest Neighbor Classification," vol. 13, no. 2, pp. 195–202, 2016.
- [3] R. I. Ndaumanu and M. R. Arief, Kusrini, "Analisis Prediksi Tingkat Pengunduran Diri Mahasiswa dengan Metode K-Nearest Neighbor," *Jatisi*, vol. 1, no. 1, pp. 1–15, 2014.
- [4] Hasmawati, Nangi, Jumadil, and M. Muchtar, "Aplikasi Prediksi Penjualan Barang Menggunakan Metode K-Nearest Neighbor (Knn) (Studi Kasus Tumaka Mart)," semanTIK, vol. 3, no. 2, pp. 151–160, 2017.
- [5] N. N. Dzikrulloh and B. D. Setiawan, "Penerapan Metode K Nearest Neighbor (KNN) dan Metode Weighted Product (WP) Dalam Penerimaan Calon Guru Dan Karyawan Tata Usaha Baru Berwawasan Teknologi (Studi Kasus: Sekolah Menengah Kejuruan Muhammadiyah 2 Kediri)," vol. 1, no. 5, pp. 378–385, 2017.
- [6] J. S. Informasi and F. Teknik, "Optimasi teknik klasifikasi modified k nearest neighbor menggunakan algoritma genetika," no. September, pp. 130–134, 2014.