

Perbandingan Metode Klasifikasi Naive Bayes Dan K-Nearest Neighbor Dalam Memprediksi Prestasi Siswa

Asih Winantu^{1*}, Chusnul Khatimah¹

¹Sistem Informasi, STMIK El Rahma, Yogyakarta, 55153, Indonesia
asihwinantu@stmikelrahma.ac.id, chusnulchusnaa@gmail.com

Abstrak

Penelitian ini bertujuan memprediksi prestasi siswa dengan cara melakukan pengolahan terhadap nilai raport siswa dengan pendekatan metode data mining klasifikasi Naïve Bayes dan *k-Nearest Neighbor*, serta memanfaatkan software Orange sebagai alat bantu pengujian. Data yang digunakan adalah data nilai siswa kelas 1 tahun ajaran 2019 sampai 2022 berjumlah 143 data. Hasil evaluasi dengan menggunakan test and score menghasilkan nilai CA (*Classification Accuracy*) algoritma Naïve Bayes 0.909 dan algoritma k-NN 0.848, nilai F1 algoritma Naïve Bayes 0.858 dan algoritma k-NN 0.820, nilai Precision algoritma Naïve Bayes 0.862 dan algoritma k-NN 0.821 nilai Recall algoritma Naïve Bayes 0.857 dan algoritma k-NN 0.820. Nilai AUC Naïve Bayes bernilai 0,909 dan k-NN bernilai 0,848.

Oleh karena itu, untuk studi kasus dievaluasikan model ROC Analysis yang memiliki nilai akurasi paling baik adalah Naïve Bayes karena kurvanya paling dekat dengan titik koordinat 0,1. Berdasarkan analisis masing-masing nilai tersebut memiliki jumlah data False Negatif dan False Positif mendekati (*Symmetric*) dan nilai akurasi sangat tinggi maka akurasi dapat digunakan sebagai acuan performansi algoritma. Dengan adanya penelitian ini diharapkan dapat mempermudah untuk menentukan potensi siswa dan dapat membantu meningkatkan sistem pembelajaran bagi siswa yang mengalami kesulitan dalam mencapai prestasi.

Kata kunci: *k-Nearest Neighbor*, Naïve Bayes, Prediksi prestasi siswa

Abstract

This study aims to predict student achievement by processing students' report cards using the Naïve Bayes classification and k-Nearest Neighbor data mining method approach, and utilizing Orange software as a testing tool. The data used is data on grade 1 student scores for the 2019 to 2022 academic year totaling 143 data. The evaluation results using test and score yield CA (Classification Accuracy) values for the Naïve Bayes algorithm 0.909 and the k-NN algorithm 0.848, the F1 value for the Naïve Bayes algorithm 0.858 and the k-NN algorithm 0.820, the Precision value for the Naïve Bayes algorithm 0.862 and the k-NN algorithm 0.821 the Recall value of the Naïve Bayes algorithm is 0.857 and the k-NN algorithm is 0.820. The Naïve Bayes AUC value is 0.909 and the k-NN value is 0.848. Therefore, for the case study, the ROC analysis model that has the best accuracy value is Naïve Bayes because the curve is closest to the coordinate point of 0.1. Based on the analysis of each of these values, the amount of False Negative and False Positive data is close to (Symmetric) and the accuracy value is very high, so accuracy can be used as a reference for algorithm performance. With this research, it is hoped that it will make it easier to determine student potential and can help improve the learning system for students who have difficulty in achieving achievement.

Keywords: *k-Nearest Neighbor*, Naïve Bayes, Prediction of student achievement

1. PENDAHULUAN

Dalam dunia pendidikan nilai menjadi salah satu penentu prestasi siswa. Prestasi belajar siswa merupakan salah satu aspek yang paling penting

dalam bidang pendidikan. Hal ini dianggap serius karena sulitnya menentukan faktor atau variabel yang mempengaruhi hasil belajar siswa. Prestasi belajar yang tinggi selalu menjadi harapan semua pihak. Bagi pihak madrasah prestasi

belajar siswanya merupakan salah satu indikator efektif proses belajar mengajar, yang sekaligus dapat digunakan untuk meningkatkan citra madrasah. Prestasi belajar yang dicapai oleh siswa berdasarkan nilai rata-rata ujian akhir mata pelajaran yang tertulis pada buku raport.

Prediksi prestasi belajar siswa dapat dijadikan sebagai dasar untuk melakukan intervensi dini terhadap potensi kegagalan siswa untuk mencapai tujuan pembelajaran dan disaat yang sama dapat melakukan perubahan strategi pembelajaran agar dapat memfasilitasi keragaman siswa.

Untuk bisa memprediksi prestasi siswa, dapat dilakukan dengan metode Data Mining. Data Mining didefinisikan sebagai sebuah proses untuk menemukan hubungan, pola dan tren baru yang bermakna dengan menyaring data yang sangat besar yang tersimpan dalam penyimpanan, menggunakan teknik pengenalan pola seperti teknik statistik dan matematika. Data Mining juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. (Marcoulides, 2005)

Tugas utama pada data mining diantaranya yaitu klasifikasi yang merupakan suatu teknik dengan melihat pada kelakuan dan atribut dari kelompok yang telah didefinisikan. Pada penelitian ini digunakan metode klasifikasi *Naive Bayes* dan *k-Nearest Neighbor*.

Algoritma K-Nearest Neighbor merupakan metode klasifikasi yang mengelompokkan data baru berdasarkan jarak data baru itu ke beberapa data/tetangga (neighbor) terdekat (Yulianti & Nurdin, 2018).

Sedangkan Algoritma Naive Bayes merupakan algoritma data mining yang dapat digunakan untuk melakukan klasifikasi, dan merupakan algoritma yang akan memberikan hasil kemungkinan yang akan terjadi di masa akan datang dari hasil klasifikasi (Devita et al., 2018).

Penelitian prediksi prestasi siswa dengan menggunakan metode Naive Bayes dan k-Nearest Neighbor, sudah sering dilakukan sebelumnya. Antara lain adalah penelitian yang dilakukan oleh Nuari et al. (2018). Penelitian ini menganalisis varietas padi yang sosok untuk lahan pertanian. Peneliti lainnya yaitu Widayati, (2021), yang melakukan penelitian dengan menggunakan data *missing value* untuk memprediksi tingkat kelulusan dari lama studi

dan IP (Indeks Predikat), menggunakan metode *Naive Bayes*.

Basri & Sari, (2020) melakukan penelitian dengan merancang sebuah sistem informasi klasifikasi hasil belajar peserta didik SDN Inpres 046 menggunakan teknik metode data mining *K-Nearest Neighbor (K-NN)* untuk memudahkan mengklasifikasi hasil belajar sekolah tersebut. Dengan sistem tersebut hasil pengklasifikasiannya dapat menjadi bahan bagi pihak sekolah dalam proses penilaian untuk mendapatkan hasil peserta didik dengan nilai terbaik setiap tahunnya

Sedangkan Suwirmayanti, (2017), melakukan penelitian untuk memberikan rekomendasi pemilihan tipe mobil menggunakan algoritma k-Nearest Neighbor,

Dengan adanya penelitian ini diharapkan dapat mempermudah untuk menentukan potensi siswa dan dapat membantu meningkatkan sistem pembelajaran bagi siswa yang mengalami kesulitan dalam mencapai prestasi.

2. METODE

2.1 Metode Pengumpulan data

Observasi yang dilakukan pada penelitian ini fokus pada buku catatan nilai siswa. Hasil yang didapat berupa catatan guru, data siswa, transkrip nilai, buku nilai atau raport. Sehingga didapatkan 143 data berupa catatan nilai akhir hasil belajar siswa kelas 1 tahun ajaran 2019 sampai 2022 dimana 10 data sebagai data validasi sehingga menyisakan 133 data yang akan digunakan sebanyak 80% sebagai data training dan 20% data testing. Data yang diperoleh akan digunakan penelitian tentunya data yang berkaitan dengan keterangan diri dari siswa dan data akademik selama masih aktif yaitu nama siswa, jenis kelamin, nilai setiap mata pelajaran, nilai hafalan, nilai praktek dan nilai kepribadian.

2.2 Preprocessing

Penelitian ini bersifat kuantitatif dengan data set yang dikumpulkan, diolah dan dianalisis untuk mencari hubungan antar variabel yang diteliti serta membandingkan dua algoritma klasifikasi sebagai penentuan algoritma terbaik. Penghitungan yang digunakan untuk memprediksi prestasi siswa MDTU Miftahul Huda dengan data mining algoritma Naive Bayes dan algoritma k-

Nearest Neighbor disertai dengan alat bantu pengujian Orange 3.

Pada tahap penelitian menggunakan KDD (*Knowledge Discovery in Database*) sebagai langkah-langkah dalam pengerjaan. Proses ini digunakan untuk menjelaskan penggalian informasi tersembunyi dalam suatu basis yang besar. Proses data mining menggunakan aplikasi microsoft excel yang merupakan aplikasi yang dapat membuat tabel.

Proses yang dilalui pada tahap KDD ini adalah :

1. Seleksi data

Data *selection* bertujuan untuk menyeleksi data dengan cara menghapus atribut yang tidak diperlukan dalam proses klasifikasi dan prediksi yaitu jumlah dan rata-rata setiap mata pelajaran. Sedangkan atribut yang dibutuhkan terdapat 14 atribut yaitu jenis kelamin, nilai UAS 6 mata pelajaran wajib, nilai UAS 4 mata pelajaran tambahan, nilai hafalan, nilai praktek dan atribut prestasi sebagai label kelas target.

2. Pembersihan Data

Pada tahap ini merupakan tahap dari *preprocessing* pada data mining yang mencakup membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (*typografi*) pada keseluruhan data. Berdasarkan hasil pembersihan data terdapat 14 atribut yang dibutuhkan berupa jenis kelamin, nilai Al-Qur'an Hadist, nilai Aqidah, nilai Fiqih, nilai Akhlaq, nilai Tarikh, nilai Bahasa Arab, nilai Diniyatul Islam, nilai Khot, nilai Imla/Ejaan nilai menggandeng, nilai hafalan, nilai praktek dan atribut prestasi sebagai label kelas target

3. Transformasi

Fase ini melakukan normalisasi atau proses mengubah atau mentransformasikan data kedalam bentuk yang paling tepat/cocok dan tidak memiliki jarak yang terlalu jauh antar data. Data di ubah dan disatukan ke dalam bentuk yang sesuai untuk dimining. Dalam data terdapat atribut kategori jenis kelamin yaitu laki-laki dan perempuan yang akan diubah kedalam bentuk diskrit menjadi bernilai 1 jika perempuan dan 2 jika laki-laki. Pengubahan atribut pada data training ini akan digunakan saat perhitungan algoritma k-Nearest Neighbor.

4. Data mining

Tahapan ini adalah tahapan yang menerapkan metode dari data mining untuk mengolah data

yang ada. Metode yang digunakan adalah Naive Bayes Classifier dan k-Nearest Neighbor dengan cara mengklasifikasi data sebenarnya dan memprediksi data berdasarkan pengujian pada orange. Nantinya akan dilakukan penghitungan berdasarkan tiap-tiap atributnya sesuai dengan kelas yang telah ditentukan, hasil yang didapat adalah prediksi data terhadap data actual serta perbandingan nilai akurasi yang tinggi bahwa algoritma tersebut merupakan metode yang cocok atau sesuai untuk digunakan berdasarkan data uji yang telah dihitung sebelumnya. Pada tahap ini akan dilakukan perhitungan algoritma Naive Bayes dengan mencari probabilitas tiap atribut dan perhitungan algoritma k-NN dengan cara mencari jarak data latih dengan data uji menggunakan rumus *euclidean distance* pada data yang telah dinormalisasi.

5. Pemodelan (*Modelling*)

Tahapan setelah dilakukannya persiapan data dilanjutkan dengan pemodelan untuk mencari, mengidentifikasi dan menampilkan pola. Pemilihan algoritma berdasarkan tipe data yang akan diklasifikasi dan diprediksi menggunakan model *Test and Score* dan *Prediction*.

3. HASIL DAN PEMBAHASAN

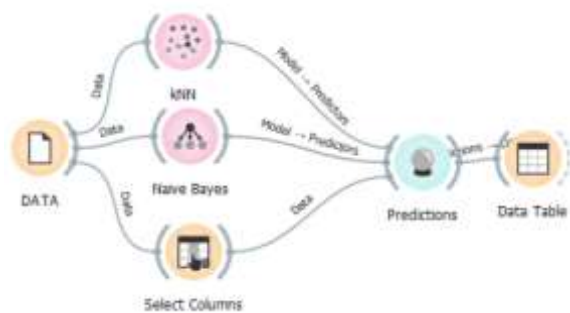
3.1 Evaluasi

Pada tahap ini akan dilakukan pemilihan kolom data dan pemisahan data training dan testing yang akan digunakan. Kolom data yang akan digunakan akan disimpan sedangkan kolom yang tidak terpakai akan dihapus. Dimana data yang digunakan adalah data siswa kelas 1 tahun ajaran 2019 sampai 2022 berjumlah 143 data dimana data akan dibagi secara proporsional berdasarkan label target yang sesuai dengan label prestasi yaitu sejumlah 10 data sebagai data validasi yang diambil dari data berlabel Ya sejumlah 6 dari 91 data dan data berlabel Tidak sejumlah 4 dari 52 data.

Data training yang digunakan sejumlah 106 data yang diambil dari data berlabel Ya sejumlah 70 dari 83 data dan data berlabel Tidak sejumlah 36 dari 50 data. Data testing yang digunakan sejumlah 27 data yang diambil dari data berlabel Ya sejumlah 13 dari 83 data dan data berlabel Tidak sejumlah 14 dari 50 data.

a. Model Prediction

Evaluasi model prediksi bertujuan untuk mengetahui kemampuan model dalam penentuan klasifikasi prestasi siswa. Dalam analisis kemampuan model, klasifikasi dalam memprediksi prestasi siswa menggunakan dua metode algoritma yaitu klasifikasi Naïve Bayes dan klasifikasi k-NN. Pengujian dilakukan dengan menerapkan 10 data validasi sebagai data yang akan diprediksi. Hasil dari data uji dibandingkan dengan data yang sebenarnya, selanjutnya dihitung akurasi dengan menggunakan persamaan *AUC/ROC*, *F1*, *CA*, *F1 Precision* dan *Recall* untuk model prediksi data.



Gambar 1. Alur model prediksi

Hasil pengujian klasifikasi dengan data uji prestasi siswa menggunakan metode Naïve Bayes dan metode k-NN dapat dilihat pada tabel 1.

Tabel 1. Hasil Uji Naïve Bayes dan k-NN

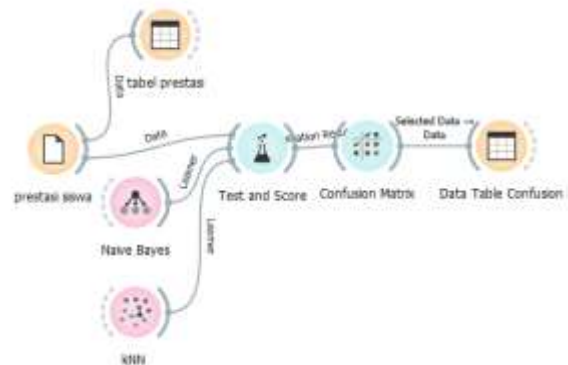
No.	Data sebenarnya	Naïve Bayes	k-NN
1	Tidak	Tidak	Tidak
2	Tidak	Tidak	Tidak
3	Tidak	Ya	Tidak
4	Tidak	Tidak	Tidak
5	YA	Ya	Ya
6	YA	Ya	Ya
7	YA	Ya	Ya
8	YA	Ya	Ya
9	YA	Ya	Ya
10	Tidak	Tidak	Tidak
Data salah	1	0	
Data Benar	9	10	
Jumlah Data	10	10	
Akurasi	90%	100%	

Pada tabel 1 tersebut menunjukkan bahwa dari 10 data validasi prestasi siswa yang diprediksi dengan menggunakan algoritma Naïve Bayes sejumlah 1 data diprediksi salah, 9 data diprediksi benar dengan menghasilkan akurasi 90%.

Sedangkan data prestasi siswa yang diprediksi menggunakan algoritma k-Nearest Neighbor sejumlah 0 data diprediksi salah, 10 data diprediksi benar dengan menghasilkan akurasi 100%. Dari hasil tersebut dapat dikatakan bahwa algoritma k-Nearest Neighbor lebih baik dalam melakukan prediksi terhadap data baru prestasi siswa.

b. Model Test and Score

Model test and score digunakan untuk mengevaluasi kinerja algoritma dalam mengklasifikasi data menggunakan metode k-Fold Cross Validation dengan bantuan *widget Test and Score*.



Gambar 2 Model klasifikasi dan Evaluasi data

Klasifikasi pada tahap ini dilakukan dengan menggunakan kumpulan data uji jika data yang digunakan 80% dari sejumlah 133 data dengan 1 atribut sebagai target, 1 atribut kategorial dan 13 atribut numerik sehingga diperoleh hasil test and score seperti terlihat pada tabel 2 berikut :

Tabel 2 Hasil Test and Score

Model	AUC	CA	F1	Precision	Recall
Naïve Bayes	0.909	0.857	0.858	0.862	0.857
kNN	0.848	0.820	0.820	0.821	0.820

Berdasarkan tabel 2, diperoleh hasil perhitungan Precision, recall, accuracy dari masing-masing model seperti terlihat pada gambar. Hasil klasifikasi model Naive Bayes dan KNN menunjukkan bahwa terdapat perbedaan nilai pada dua model AUC, diketahui bahwa nilai AUC yang paling tinggi adalah

metode Naive Bayes yaitu 0.909. AUC dipakai untuk mengukur kinerja diskriminatif dengan memperkirakan probabilitas output dari ilustrasi yang diseleksi secara acak dari populasi positif ataupun negatif, semakin besar AUC, semakin baik hasil klasifikasi yang digunakan.

c. Confusion Matrix

		Predicted		Σ
		TIDAK	YA	
Actual	TIDAK	43	7	50
	YA	12	71	83
Σ		55	78	133

Gambar 3 Confusion Matrix Naïve Bayes

Pada gambar 1 tersebut menunjukkan bahwa nilai dari True Positif (TP) adalah 71, True Negatif (TN) adalah 43, False Positif (FP) adalah 7, dan False Negatif (FN) adalah 12. Maka nilai Accuracy, Precision dan Recall dari metode Naïve Bayes adalah sebagai berikut:

$$Accuracy = \frac{(71+43)}{(71+43+7+12)} = 0.86 = 86\%$$

$$Precision = \frac{(71)}{(71+7)} = 0.91 = 91\%$$

$$Recall = \frac{(71)}{(71+12)} = 0.85 = 85\%$$

$$F1 = \frac{(2 \cdot 0.85 \cdot 0.91)}{(0.85 + 0.91)} = 0.88 = 88\%$$

		Predicted		Σ
		TIDAK	YA	
Actual	TIDAK	39	11	50
	YA	13	70	83
Σ		52	81	133

Gambar 4 Confusion Matrix k-Nearest Neighbor

Pada Gambar 4 menunjukkan bahwa nilai dari True Positif (TP) adalah 70, True Negatif (TN) adalah 39, False Positif (FP) adalah 11, dan False Negatif (FN) adalah 13. Maka nilai Accuracy, Precision dan Recall dari metode KNN adalah sebagai berikut:

$$Accuracy = \frac{(70+39)}{(70+39+11+13)} = 0.82 = 82\%$$

$$Precision = \frac{(70)}{(70+11)} = 0.86 = 86\%$$

$$Recall = \frac{(70)}{(70+13)} = 0.84 = 84\%$$

$$F1 = \frac{(2 \cdot 0.86 \cdot 0.84)}{(0.86 + 0.84)} = 0.85 = 85\%$$

Berdasarkan hasil evaluasi dan validasi dengan menggunakan Confusion Matrix diperoleh nilai perbandingan Accuracy, Precision dan Recall dari 2 metode Naïve Bayes dan KNN seperti terlihat pada Tabel 2.

Tabel 3 Perbandingan Kinerja

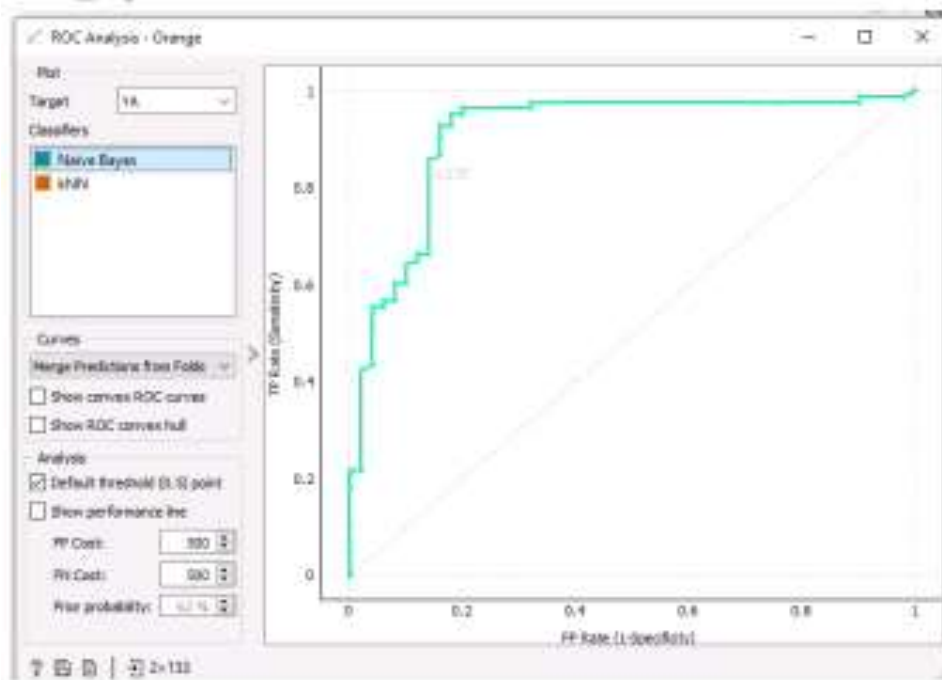
Model	Accuracy	Precision	Recall	F1
Naïve Bayes	86%	91%	85%	88%
kNN	82%	86%	84%	85%

Berdasarkan tabel 3 dapat diketahui bahwa kinerja dari model Naïve Bayes lebih baik dari model KNN. Akurasi klasifikasi tidak bisa mencapai hasil yang sempurna karena pasti ada nilai eror. Hal tersebut dipengaruhi oleh banyaknya data uji dan data latih yang digunakan dalam proses simulasi maupun normalisasi yang dilakukan.

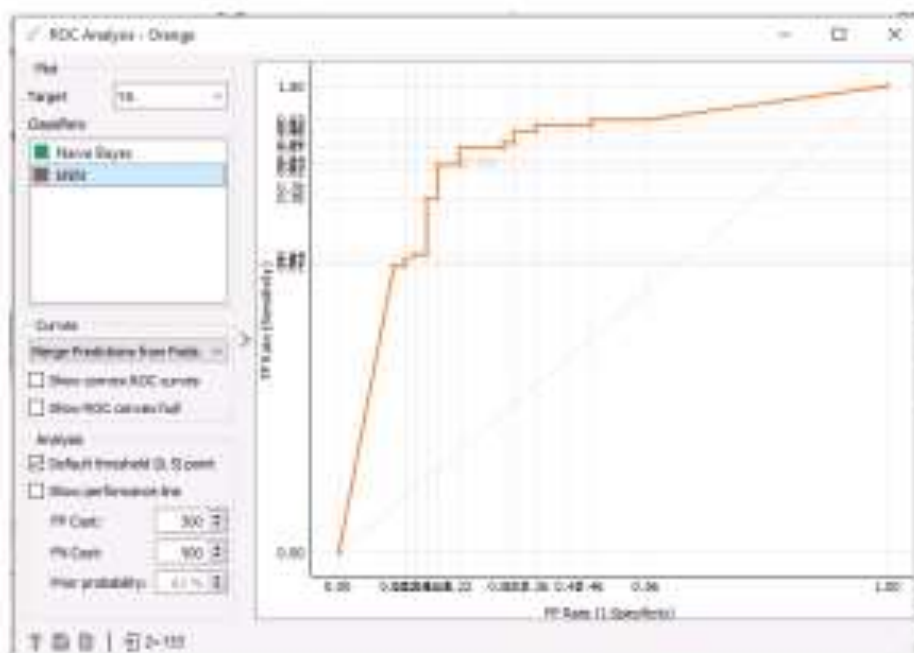
d. Evaluasi Model ROC Analysis

Kurva ROC adalah suatu kurva yang terbentuk dari titik koordinat antarsensitivitas yang berada di sumbu Y dan 1-spesifisitas yang berada di sumbu X. Setiap titik koordinat mewakili sensitivitas dan 1-spesifisitas yang dihasilkan oleh tiap nilai hasil pengukuran confusion matrix. Penilaian terhadap kemampuan suatu test dilakukan dengan menggunakan luas area *under the curve* (AUC).

AUC meliputi keseluruhan area di bawah kurva yang terbentuk dari semua koordinat sensitivitas dan 1-spesifisitas. Nilai AUC berkisar dari 0 – 1, semakin luas AUC maka semakin baik kemampuan suatu test untuk mendeteksi suatu penyakit. Kemampuan suatu test dinyatakan baik jika $AUC \geq 0,7$.



Gambar 5 Kurva ROC Naïve Bayes



Gambar 6 Kurva ROC k-NN

Nilai akurasi secara manual bisa dilakukan dengan melihat perbandingan curve ROC yang divisualisasi dari Confusion Matrix. Model melihat kurva ROC adalah cara yang paling mudah terlihat untuk membandingkan nilai akurasi masing-masing model klasifikasi secara grafis. Semakin tinggi True Positive Rate dan semakin kecil False Positive Rate maka thresholdnya semakin bagus.

Hasil grafis ROC dapat dilihat pada gambar 6 menunjukkan bahwa hasil analisis model perbandingan dengan AUC Naïve Bayes bernilai 0,909 dan k-NN bernilai 0,848. Oleh karena itu, untuk studi kasus dievaluasi model ROC Analysis yang memiliki nilai akurasi paling baik adalah Naïve Bayes karena kurvanya paling dekat dengan titik 0,1.

4. KESIMPULAN

Data yang digunakan untuk klasifikasi adalah data prestasi siswa dengan fitur nama, jenis kelamin, nilai Al-Qur'an Hadist, nilai Aqidah, nilai Fiqih, nilai Akhlaq, nilai Tarikh, nilai Bahasa Arab, nilai Diniyatul Islam, nilai Khot, nilai Im-lah/Ejaan nilai menggandeng, nilai hafalan, nilai praktek dan atribut prestasi sebagai label kelas target.

Hasil evaluasi dengan menggunakan test and score menghasilkan nilai CA (*Classification Accuracy*) algoritma Naïve Bayes 0.909 dan algoritma k-NN 0.848, nilai F1 algoritma Naïve Bayes 0.858 dan algoritma k-NN 0.820, nilai Precision algoritma Naïve Bayes 0.862 dan algoritma k-NN 0.821 nilai Recall algoritma Naïve Bayes 0.857 dan algoritma k-NN 0.820. Berdasarkan analisis masing-masing nilai tersebut memiliki jumlah data False Negatif dan False Positif mendekati (*Symmetric*) dan nilai akurasi sangat tinggi maka akurasi dapat digunakan sebagai acuan performansi algoritma.

Hasil grafis ROC Analysis dapat dilihat pada gambar 6 menunjukkan bahwa hasil analisis model perbandingan dengan AUC Naïve Bayes bernilai 0,909 dan k-NN bernilai 0,848. Oleh karena itu, untuk studi kasus dievaluasi model ROC Analysis yang memiliki nilai akurasi paling baik adalah Naïve Bayes karena kurvanya paling dekat dengan titik koordinat 0,1.

Hasil evaluasi kemampuan prediksi terhadap 10 data nilai siswa sejumlah 1 data diprediksi salah, 9 data diprediksi benar dengan menghasilkan akurasi 90%. Sedangkan data prestasi siswa yang diprediksi menggunakan algoritma K-Nearest Neighbor 10 data diprediksi benar dengan menghasilkan akurasi 100%. Dari hasil tersebut dapat dikatakan bahwa algoritma k-Nearest Neighbor lebih baik dalam melakukan prediksi terhadap data prestasi siswa.

Metode Naive Bayes Dan K-Nearest Neighbor Untuk Klasifikasi Artikel Berbahasa Indonesia. *JTIK*, 5(4), 427–434.

Marcoulides, G. A. (2005). *Discovering Knowledge in Data: an Introduction to Data Mining*. *Journal of the American Statistical Association*, 100(472).

Nuari, R., Apriliyani, A., Juwari, & Kusri. (2018). Implementasi Metode K-Nearest Neighbor (Knn) Untuk Memprediksi Varietas Padi Yang Cocok Untuk Lahan Pertanian. *Informa*, 4(2), 2–8.

Pratiwi, L. (2020). Clustering Data Kelulusan Mahasiswa STMIK EL-Rahma Menggunakan Metode K-Means. *STMIK El Rahma Yogyakarta*.

Suwirmayanti, N. L. G. P. (2017). Penerapan Metode K-Nearest Neighbor Untuk Sistem Rekomendasi Pemilihan Mobil. *Techno*, 16(2), 120–131.

Widayati, T. (2021). Pengaruh Data Missing Value Pada Prediksi Kelulusan Mahasiswa STMIK El Rahma Yogyakarta Menggunakan Metode Klasifikasi Naïve Bayes. *STMIK El Rahma Yogyakarta*.

Yulianti, E., & Nurdin, Y. A. (2018). Sistem Pendukung Keputusan Penerimaan Bantuan Siswa Miskin (BSM) Berbasis Online Dengan Metode KNN (K-Nearest Neighbor) (Studi kasus : SMPN 1 Koto XI Tarusan). *Teknoif*, 6(1), 12–17.

DAFTAR PUSTAKA

Basri, A. F., & Sari, C. R. (2020). Sistem Informasi Pengklasifikasian Hasil Belajar Peserta Didik Dengan Teknik Data Mining Metode K-Nearest Neighbor (K-NN). *Journal Pegguruang: Conference Series*, 2(1), 79–84.

Devita, R. N., Herwanto, H. W., & Wibawa, A. P. (2018). Perbandingan Kinerja