

Analisis Tingkat Kepuasan Siswa Dalam Pembelajaran *Hybrid* Menggunakan Algoritma *K – Nearest Neighbor* (KNN)

Student Satisfaction Level Analysis In Hybrid Learning Using K – Nearest Neighbor (KNN) Algorithm

Rizky Irfan Sulthoni¹, Luthfi Ali Muharom^{2*}, Miftahur Rahman³

¹Fakultas Teknik, Universitas Muhammadiyah Jember

Email: rizkyirfansulthoni@gmail.com

²Fakultas Teknik, Universitas Muhammadiyah Jember *Koresponden Author

Email: luthfialimuharom@unmuhjember.ac.id

³Fakultas Teknik, Universitas Muhammadiyah Jember

Email: miftahurrahman@unmuhjember.ac.id

Abstrak

Virus *Corona* sudah menyerang di berbagai penjuru dunia, tidak terkecuali di Indonesia. MtsN 11 Jember adalah salah satu sekolah yang menerapkan sistem pembelajaran yang membagi kelas menjadi sebagian daring dan luring atau *hybrid*. Kenyamanan belajar siswa menentukan penyerapan ilmu pengetahuan maka dibutuhkan survey untuk menentukan tingkat kepuasan siswa selama pembelajaran *hybrid*, guna untuk mempermudah pihak sekolah dalam perencanaan pembelajaran kedepannya digunakan teknik klasifikasi data mining dan algoritma *K – NEAREST NEIGHBOR* menggunakan *rapidminer*. Hasil dari penelitian ini adalah prediksi tingkat kepuasan siswa selama pembelajaran *hybrid*. berdasarkan rata - rata nilai akurasi 86.78%, presisi 40.00%, dan *recall* 24.00% terhadap klasifikasi tingkat kepuasan pembelajaran siswa.

Kata Kunci : Survey, Data Mining, *K – NEAREST NEIGHBOR* (KNN)

Abstract

The Corona virus has attacked all over the world, Indonesia is no exception. MtsN 11 Jember is one of the schools that implements a learning system that divides the class into online and offline or hybrid parts. The convenience of student learning determines the absorption of knowledge, so a survey is needed to determine the level of student satisfaction during hybrid learning, in order to facilitate the school in planning future learning using data mining classification techniques and the *K – NEAREST NEIGHBOR* algorithm using *rapidminer*. The results of this study are predictions of the level of student satisfaction during hybrid learning, based on the average value of accuracy 86.78%, precision 40.00%, and recall 24.00 % on the classification of student learning satisfaction levels.

Keywords : Survey, Data Mining, *K – NEAREST NEIGHBOR* (KNN)

1. PENDAHULUAN

Virus *Corona* sudah menyerang diberbagai penjuru dunia, tidak terkecuali di Indonesia. Satu tahun lebih warga Indonesia terpapar dari Virus *Corona*. Banyak orang yang meyakini virus ini berasal dari Wuhan, China. Sejumlah negara mewajibkan warganya seperti menjaga jarak, menjauhi kerumunan, memakai masker dan lain sebagainya

Kebijakan terbaru pemerintah yaitu PPKM Level 1 – 4 yang dimana mengizinkan sekolah mengadakan pembelajaran tatap muka yang dilakukan secara ketat / dibatasi 50% siswa. 50% siswa melakukan pembelajaran tatap muka dan 50% siswa lainnya melakukan pembelajaran secara daring / online. Istilah tersebut adalah pembelajaran *hybrid*. Banyak pendapat siswa yang setuju dan tidak setuju terhadap pembelajaran *hybrid* tersebut. Siswa yang setuju kebanyakan merasa nyaman melakukan pembelajaran *hybrid* karena fleksibilitasnya dan siswa yang tidak setuju kebanyakan karena terkendala dengan fasilitas yang ada. [1]

K – Nearest Neighbor merupakan algoritma yang digunakan untuk memproses klasifikasi hasil survey. Kelebihan dari algoritma *K – Nearest Neighbor* terbukti dapat mencapai akurasi presisi dan *recall* yang tinggi. Berdasarkan penelitian dari penentuan penerima bantuan sosial bagi siswa yang terkena dampak Covid-19 menggunakan metode *K – Nearest Neighbor* (KNN) oleh Mochammad Faris didapatkan hasil, akurasi sebesar 98%, presisi 100%, dan *recall* 96% dari 50 data yang digunakan dibagi menjadi 40 data training dan 10 data testing [2] Berdasarkan dampak - dampak yang ada maka fokus yang akan diteliti adalah seluruh siswa dengan menggunakan metode survey. Metode survey yang digunakan adalah memberikan *google form* kepada seluruh siswa. Setelah survey selesai kemudian diklasifikasi menggunakan metode *K – Nearest Neighbor* yang merupakan proses untuk mengetahui tentang pandangan seluruh siswa terhadap peristiwa yang terjadi, baik itu berupa pandangan setuju ataupun pandangan yang tidak setuju. Kemudian hasil klasifikasi yang didapat diberikan kepada pihak

sekolah untuk dijadikan pertimbangan untuk pembelajaran selanjutnya.

2. KAJIAN LITERATUR

Penelitian ini menggunakan beberapa landasan teori di dalamnya. Berikut beberapa teori-teori penting yang digunakan dalam penelitian ini :

A. Pengumpulan Data Kuisoner

Kuisoner adalah teknik pengumpulan data yang dilakukan dengan cara memberi seperangkat pertanyaan tertulis kepada responden untuk dijawabnya [3]

Kuisoner yang digunakan peneliti yaitu melalui *google form* yang ditujukan kepada siswa MTsN 11 Jember untuk mengetahui seberapa puas dan tidak puas siswa terhadap proses pembelajaran *hybrid*.

B. Dimensi Kepuasan Siswa

Kepuasan siswa dalam pembelajaran *hybrid* merupakan perasaan siswa antara harapan dan kenyataan. Ada banyak indikator tentang kepuasan siswa selama pandemi covid-19 ini. Tolak ukur yang menjadi kepuasan siswa dapat dilihat dalam hal profesionalisme guru dan kesempatan belajar. [4] mengemukakan ada lima dimensi dimensi Serqual untuk kepuasan belajar siswa, yaitu:

1. *Tangible* (Kualitas Pelayanan), Hal ini mengacu pada aspek fisik sekolah yang diperlukan untuk mendukung pembelajaran. Suatu bukti menjadi signifikan untuk kemampuan kita dalam menyediakan layanan nyata dan perlengkapan sekolah. Secara khusus, memberikan peluang digital untuk pembelajaran *online* melalui pembelajaran virtual.
2. *Assurance* (Kepastian), Lingkungan sekolah memberikan jaminan kualitas perilaku guru kepada siswa untuk menanamkan rasa percaya diri dalam proses belajar mengajar, termasuk kompetensi profesional guru. Selama pandemi covid-19 ini, sekolah memberikan pembelajaran *online* melalui pembelajaran virtual, sehingga guru membutuhkan

persiapan untuk membantu siswa yang mengalami kesulitan belajar.

3. Empati, Dalam hal ini, guru memberikan contoh moral bagi siswanya melalui tindakan dan komunikasi dengan orang lain. Perhatian dari seorang guru terhadap siswa merupakan layanan yang baik. Guru tidak hanya menjadi seorang pendidik, tetapi juga sebagai penasihat seorang siswa.
4. *Reliability* (Keandalan), Sekolah seharusnya menyediakan fasilitas dan layanan yang memadai. Keandalan juga berkaitan dengan kemampuan guru dalam mengkoordinasikan dan mengembangkan kurikulum, serta penggunaan alat pembelajaran melalui penggunaan *WhatsApp*, *Google Classroom*, *Zoom Meeting* dan *Google Meeting* secara efektif.
5. *Responsiveness* (daya tanggap), Kesiapan sekolah untuk mendengarkan, menanggapi, dan memecahkan masalah bagi siswa yang berkaitan dengan pembelajaran, pengajaran atau masalah pribadi yang mempengaruhi konsentrasi.

C. *Data mining*

Data mining juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data [5]. Beberapa tahapan dalam proses *data mining* adalah :

1. *Data cleaning*, merupakan proses menghilangkan *noise* dan data yang tidak konsisten atau data tidak relevan dan dianggap mengganggu pada data yang akan diproses.
2. Integrasi data, merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru atau data utuh.
3. Seleksi data, merupakan pemilihan data yang tepat dari sekumpulan data yang telah diperoleh. Hal ini berfungsi agar data yang diteliti dengan topik yang diambil dapat selaras.
4. Transformasi data, merupakan perubahan dan penggabungan data ke dalam data yang

dijadikan format pada penelitian yang dilakukan.

5. Proses *mining*, merupakan implementasi metode atau algoritma ke dalam data untuk memperoleh model yang diharapkan.
6. Evaluasi data, merupakan langkah mengidentifikasi hasil pola yang diperoleh untuk menentukan tingkat keberhasilan penelitian yang dilakukan.

D. *K-Nearest Neighbor (KNN)*

K-Nearest Neighbor (KNN) merupakan metode yang biasa digunakan pada klasifikasi data. Algoritma ini digunakan untuk mengklasifikasikan terhadap objek berdasarkan data pembelajaran yang jarak tetangganya paling dekat atau memiliki nilai selisih yang kecil dengan objek tersebut. [6]

$$\left(\sum_{k=1}^m (x_{ik} - x_{jk})^2 \right)^{1/2}$$

Dimana, O adalah nilai observasi atau pengamatan, E adalah nilai *expected* atau harapan, B adalah \sum baris dan K adalah \sum kolom.

Dimana, X_{ik} = nilai X pada training data, X_{jk} = nilai X pada testing data, m = batas jumlah banyaknya data Langkah-langkah untuk menghitung algoritma k-NN:

Menentukan nilai k . Menghitung kuadrat jarak euclid (*query instance*) masing-masing objek terhadap *training* data yang diberikan. Kemudian mengurutkan objek - objek tersebut ke dalam kelompok yang mempunyai jarak euclid terkecil. Mengumpulkan label class Y (klasifikasi *Nearest Neighborhood*). Dengan menggunakan kategori *Nearest Neighborhood* yang paling mayoritas maka dapat diprediksikan nilai *query instance* yang telah dihitung.[7]

E. Akurasi , Presisi dan Recall

Performa dari suatu model kasifikasi dapat diukur dengan tingkat akurasi berdasarkan *Confusion matrix* [8].*Confusion Matrix* adalah teknik yang digunakan sebagai evaluasi model klasifikasi untuk memperkirakan objek yang benar atau salah. Sebuah *matrix* dari prediksi yang akan dibandingkan dengan kelas sebenarnya atau dengan kata lain berisi informasi nilai sebenarnya dan prediksi pada klasifikasi [5].

		Predictive Class	
		FALSE	TRUE
ACTUAL Class	FALSE	TN	FP
	TRUE	FN	TP
Total		P'	N'

Gambar 1. Confusion Matrix 2 kelas
 Sumber (Anwar, 2017)

Akurasi merupakan persentase dari data yang diprediksi secara benar. Presisi Merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. *Recall* Merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Berikut perhitungan akurasi, Presisi, *recall* adalah : [8]

$$Akurasi = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

$$Presisi = \frac{(TP)}{(FP + TP)}$$

$$Recall = \frac{(TP)}{(FN + TP)}$$

TP : *True positives*, merupakan jumlah data dengan kelas positif yang diklasifikasikan positif. TN : *True negatives*, merupakan jumlah data dengan kelas negatif yang diklasifikasikan negatif. FP : *False positives*, merupakan jumlah data dengan kelas positif diklasifikasikan negatif. FN : *False negatives*, merupakan jumlah data dengan kelas negatif diklasifikasikan positif.

F. Klasifikasi

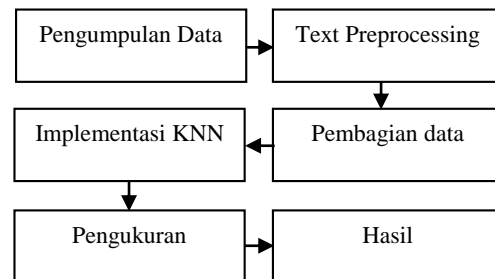
klasifikasi merupakan suatu teknik *Data mining* yang melihat sifat dari atribut dari kelompok data yang telah didefinisikan. Teknik ini dapat digunakan untuk memberi pengetahuan pada data baru dengan memanipulasi data yang ada yang telah diklasifikasikan dan dengan menggunakan hasilnya untuk memberikan pengetahuan atau sejumlah aturan. Aturan tersebut digunakan data baru untuk dapat diklasifikasikan terhadap suatu kategori atau kelas tertentu. [5].

G. Rapid Miner

Rapid Miner merupakan perangkat lunak yang bersifat terbuka (*open source*). *Rapid Miner* adalah sebuah solusi untuk melakukan analisis terhadap *Data mining*, *text mining* dan analisis prediksi. *Rapid Miner* menggunakan

berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. *Rapid Miner* memiliki kurang lebih 500 operator *Data mining*, [5]

3. METODE PENELITIAN



Gambar 2. Tahapan Penelitian
 Sumber Hasil Penelitian

A. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data survey di MTsN 11 Jember yang diambil selama 14 hari dimulai pada tanggal 21 Maret 2022 sampai 4 April 2022.

B. Text preprocessing

Pada tahap *text preprocessing* ini teknik yang digunakan adalah data *Cleaning*, yaitu menghapus data yang tidak cocok / tidak relevan.

C. Pembagian Data

Total data pada penelitian ini adalah 128 data hasil survey. Data akan dibagi menjadi data latih dan data uji dengan 5 kali percobaan. Data latih sebesar 90% sebesar 107 data, 80% sebesar 95 data, 70% sebesar 83 data, 60% sebesar 71 data, 50% sebesar 59 data. Data uji 50% sebesar 60 data, 40% sebesar 48 data, 30% sebesar 36 data, 20% sebesar 24 data, 10% 12 data.

D. Implementasi KNN

Dalam implementasi *KNN* dibagi menjadi 5 percobaan dari percobaan yang dilakukan menggunakan *rapid miner*.

E. Pengukuran

Pengukuran pada tahap ini difokuskan pada uji akurasi, presisi dan *recall*.

F. Hasil

Hasil yang akan dicatat pada penelitian ini adalah tingkat akurasi, presisi dan *recall* pada setiap percobaan yang dilakukan.

4. HASIL PENELITIAN DAN PEMBAHASAN

Pada proses ini akan dibagi menjadi tiga yaitu pengolahan data survey dan pengujian KNN dan hasil klasifikasi.

A. Pengolahan data survei

Pengolahan data survey didapat adalah 128 data. Kemudian data tersebut dilakukan proses *cleaning* dan menjadi 119 data.

B. Pengujian KNN

Pada pengujian KNN dilakukan pada 5 percobaan. Berikut hasil pengujian pada metode KNN di bawah ini dapat dijelaskan yaitu, hasil pengujian data klasifikasi menggunakan metode KNN memperoleh nilai rata – rata akurasi 86.78%, presisi 40.00% dan *recall* 24.00%

Tabel 1. Hasil pemodelan pada metode KNN

Percobaan Ke -	Akurasi (%)	Presisi (%)	<i>Recall</i> (%)
1	100.00%	100.00%	100.00%
2	83.33%	100.00%	20.00%
3	80.56%	0.00%	0.00%
4	83.33%	0.00%	0.00%
5	86.67%	0.00%	0.00%
Rata - Rata	86.78%	40.00%	24.00%

Sumber: Hasil Perhitungan

Pemodelan pada metode KNN yang diparkan pada tabel 1. di atas, diketahui bahwa model dengan akurasi presisi dan *recall* terbaik yaitu pada percobaan pertama yaitu 100%

C. Hasil klasifikasi

Pada langkah terakhir adalah memberikan hasil klasifikasi kepada pihak sekolah agar bisa digunakan sebagai pertimbangan bagaimana pendapat siswa tentang pembelajaran hybrid

5. HASIL DAN SARAN

A. Kesimpulan

Dari serangkaian proses yang dilakukan pada penelitian ini, berikut poin-poin yang dapat disimpulkan:

1. Tingkat akurasi, presisi dan *recall* dari pengklasifikasian pada pengujian algoritma KNN dapat diketahui dengan proses *performance testing* di *rapidminer*. Adapun nilai hasil rata-rata akurasi 86.78%, presisi 40.00%, dan *recall* 24.00%.
2. penerapan algoritma *K – Nearest Neighbor* (KNN) dan pengimplementasiannya dengan cara melakukan lima kali percobaan dari data hasil survey yang dimana hasil dari algoritma *K – Nearest Neighbor* (KNN) yaitu siswa setuju dengan pembelajaran *hybird*.
3. saran keputusan terhadap algoritma *K – Nearest Neighbor* (KNN) dengan cara memberikan hasil penelitian terhadap pihak sekolah yang kemudian dijadikan pertimbangan pihak sekolah untuk pembelajaran kedepannya.

B. Saran

Untuk pengembangan penelitian tentang Klasifikasi Tingkat kepuasan siswa dalam pembelajaran *hybird* menggunakan algoritma *K – Nearest Neighbor* (KNN), penulis memberikan saran yaitu:

1. Pada penelitian ini, penulis hanya menggunakan satu sekolah sebagai acuan pengambilan data survey yang dimana diharapkan pengembang dapat menambahkan beberapa sekolah agar lebih banyak data yang dapat diolah.
2. Pada penelitian ini, hanya menggunakan algoritma *K – Nearest Neighbor* sebagai acuan penelitian ini. Diharapkan pengembang dapat menambahkan algoritma yang berbeda seperti *naïve bayes*, *decision tree* sebagai perbandingan sudut pandang terhadap penelitian selanjutnya.

6. DAFTAR PUSTAKA

- [1] (Hidayat & Andira, 2019) Pengaruh Model Pembelajaran *Hybrid Learning* Berbantuan Media *Schoology* Terhadap Hasil Belajar Peserta Didik Kelas XI MIA MAN Pangkep
- [2] (Faris M, 2021). Penentuan Penerima Bantuan Sosial Bagi Siswa Yang Terkena Dampak Covid-19 Menggunakan Metode *K-Nearest Neighbor*
- [3] sugiyono (2017:35) Sugiyono. 2017. Metode Penelitian Kuantitatif, Kualitatif dan R&D. Alfabeta Bandung.
- [4] Saefani, I. (2019). Analisis Dimensi Kualitas Pelayanan Terhadap Kepuasan Mahasiswa Pada Layanan Bidang Kemahasiswaan.
- [5] (Mochamad Rizki Ilham, 2016) Implementasi *Data Mining* Menggunakan Algoritma C4.5 Untuk Prediksi Kepuasan Pelanggan Taksi Kosti.
- [6] (Rasepta, 2015) Klasifikasi Status Gizi Balita Menggunakan Metode *Modified K-Nearest Neighbor*. Universitas Islam Negeri Sultan Syarif Kasim Riau Pekanbaru.
- [7] (Hutami & Astuti, 2016) Implementasi Metode *K-Nearest Neighbor* Untuk Prediksi Penjualan Furniture Pada CV.Octo Agung Jepara.
- [8] (Anwar, 2017) Penggunaan Data Mining Untuk Memprediksi Trending Topik Menggunakan Algoritma *K-Nearest Neighbor* dan *Naive Bayes*
- [9] (Habibie, 2021) Penerapan alogaritma *K-Nearest Neighbor* untuk klasifikasi penyakit liver.