

DETEKSI DINI ASD (*AUTISM SPECTRUM DISORDER*) MENGGUNAKAN MACHINE LEARNING

Taftazani Ghazi Pratama^{a,*}, Achmad Ridwan^b, Agung Prihandono^c

^aIlmu Komputer, Universitas Muhammadiyah Kudus

^bSistem Informasi, Universitas Muhammadiyah Kudus

Purwosari, Kudus, Indonesia

Email:taftazanighazi@umkudus.ac.id

Abstrak

Deteksi dini ASD pada seorang balita sangat membantu orang tua untuk mengetahui kembang tumbuh anak. Pada penelitian ini bertujuan untuk deteksi dini ASD menggunakan *Naive Bayes* dan KNN yang diterapkan pada dataset *Autism screening data for toddlers*. Penelitian ini dilakukan melalui tahapan pra pengolahan, pembagian data training 80% dan testing 20%, pembuatan model, dan evaluasi dari model yang dibuat. Hasil evaluasi model yang dibuat menunjukkan bahwa KNN memperoleh nilai akurasi, sensitivitas, dan spesifisitas lebih tinggi daripada *Naive Bayes*. KNN memperoleh nilai akurasi sebesar 95,73%, sensitivitas sebesar 93,84%, dan spesifisitas 100%. Hal ini mengindikasikan bahwa KNN dapat membantu dalam deteksi dini pada seorang balita dengan kinerja yang baik.

Kata Kunci: deteksi dini, ASD, *Naive Bayes*, KNN

Abstract

Early detection of ASD in a toddler really helps parents to find out their child's growth and development. In this study the aim was for early detection of ASD using Naive Bayes and KNN which were applied to the Autism screening data for toddlers dataset. This research was conducted through pre-processing stages, distribution of 80% training data and 20% testing, model creation, and evaluation of the model made. The results of the evaluation of the model made show that KNN obtained higher accuracy, sensitivity and specificity values than Naive Bayes. KNN obtained an accuracy of 95.73%, a sensitivity of 93.84%, and a specificity of 100%. This indicates that KNN can assist in early detection of a toddler with good performance.

Keywords: early detection, ASD, *Naive Bayes*, KNN

I. PENDAHULUAN

Menurut data yang dihimpun oleh WHO, sekitar 1 dari 100 anak-anak di dunia ini mengalami ASD (*Autism Spectrum Disorder*) (WHO, 2023). ASD adalah sekumpulan gangguan sosial, komunikasi yang bermasalah, dan perilaku yang berulang dan terbatas dapat diidap seorang individu sejak balita hingga dewasa. Anak yang mengidap ASD dalam proses pertumbuhan dan perkembangan fisik memiliki perbedaan cara dibandingkan dengan anak-anak normal lainnya. Hal tersebut disebabkan oleh adanya penurunan kemampuan kognisi secara bertahap (Katilik & Djie, 2022). Menurut

CDC(*Centers for Disease Control and Prevention*) seorang individu bisa diketahui apakah mengalami ASD atau tidak dengan melalui 2 tahapan: 1. deteksi dini 2. Wawancara seorang individu dengan seorang psikiater, dokter, dan ahli saraf. ASD bisa dideteksi ketika seorang individu masih berumur kurang lebih 18 bulan. Dalam deteksi dini ASD saat ini menggunakan instrumen Autism Behavior Checklist (ABC), Autism Spectrum Quotient

(AQ), Child Behavior Checklist (CBCL), dan lain-lain (Rishel et al., 2005). Dalam instrumen tersebut terdapat banyak pertanyaan yang berkaitan dengan

karakteristik ASD. Pertanyaan-pertanyaan yang ada dapat diisi oleh orang tua, guru, perawat, dan pasien itu sendiri. Namun dengan banyaknya pertanyaan yang harus ini dapat membutuhkan waktu yang lama dan tidak efisien. Oleh karena itu dibutuhkan suatu teknologi yang dapat mempercepat dan tetap menghasilkan keakuratan yang baik. Dalam beberapa tahun terakhir ini proses screening, diagnosis maupun pengambilan keputusan dalam bidang biomedis maupun kedokteran dibantu dengan teknologi yang bernama *machine learning*. Beberapa penelitian berkaitan dengan *machine learning* yang digunakan dalam bidang medis maupun non medis telah dilakukan. Penelitian yang dilakukan oleh Ade Riani dkk tentang prediksi penyakit jantung menggunakan naive bayes menghasilkan nilai akurasi sebesar 86% (Riani et al., 2019) Penyakit Ginjal Kronis dapat dideteksi menggunakan metode naive bayes. Kinerja model yang diperoleh antara lain akurasi sebesar 96,43%, recall sebesar 93,18% dan precision sebesar 93,02% (A'yuniyah et al., 2022). Selain di bidang medis, *machine learning* juga dapat digunakan pada bidang yang lain. Penelitian tentang klasifikasi tingkat kelulusan siswa menggunakan KNN(K-Nearest Neighbor) memperoleh hasil akurasi sebesar 96,49% (Purwaningsih & Nurelasari, 2021). Prediksi kelompok UKT menggunakan KNN guna mengetahui masuk ke kolompok UKT yang mana didapatkan hasil akurasi yang cukup baik yaitu sebesar 84,21% (Sukamto et al., 2020). Berdasarkan penelitian-penelitian sebelumnya bahwa algoritma Naive Bayes dan KNN dapat menghasilkan tingkat keakuratan yang cukup bagus. Oleh karena itu pada penelitian ini akan menggunakan algoritma Naive bayes dan KNN. Tujuan dari penelitian ini adalah Mendeteksi dini ASD menggunakan Algoritma Naive bayes dan KNN dan mengukur kinerja model machine learning yang dibuat.

II. LANDASAN TEORI

A. Data Mining

Data mining secara umum merupakan Konsep dan Teknik untuk mengolah data atau informasi yang akan digunakan dalam

berbagai aplikasi. Secara khususnya, ini menjelaskan penyimpulan data dan alat yang digunakan untuk mengekstrak informasi untuk diambil pola dari data yang disimpan . Alat yang digunakan untuk mengekstrak informasi dari data yang disimpan. Istilah ini dikenal sebagai kunci penemuan pola pengetahuan data . konsep ini berfokus pada validitas, penggunaan, efektivitas, dan skalabilitas dari teknik penggalian data berskala besar atau kumpulan data besar. Kemudian, metode-metode tersebut yang terlibat dalam menambang pola data, asosiasi data, dan korelasi data sering digunakan dalam kumpulan data besar . konsep-konsep ini merinci metode untuk klasifikasi data dan memperkenalkan konsep dan metode untuk pengelompokan data (Wu et al., 2008).

B. Naïve Bayes

Naïve Bayes secara umum merupakan teknik penggalian data atau data mining dengan menggunakan pendekatan teori-teori probabilitas untuk membuat sebuah model klasifikasi berdasarkan pada kejadian sudah lalu yang mempunyai potensi membuat sebuah objek baru yang dikategorikan sebagai kelas yang memiliki probabilitas terbaik dan teruji.

Naïve Bayes dipilih karena berkemampuan yang cepat dalam membuat model, mempunyai banyak kemampuan memprediksi dan juga menyediakan metode terbaru dalam menggali dan memahami data. Algoritma Naïve Bayes hanya mendukung pada atribut yang bertipe data discrete atau discretized, atau tidak mendukung atribut yang bernilai continuous (numerik) dan semua atribut dapat menjadi independen, menjadi atribut yang memberi kontribusi kepada atribut yang diprediksi(Ridwan et al., 2018). Secara umum *Naïve Bayes* mempunyai persamaan :

$$P(H | X) = \frac{P(X | H) P(H)}{P(X)} \quad (1)$$

X = Data dengan class yang belum diketahui
H = Hipotesis data X merupakan suatu class spesifik

$P(H|X)$ = Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob.)

$P(H)$ = Probabilitas hipotesis H (prior prob.)

$P(X|H)$ = Probabilitas X berdasarkan kondisi tersebut

$P(X)$ = Probabilitas dari X

(2)

C. K-Nearest Neighbor

K-Nearest Neighbor merupakan paling sederhana dari algoritma *Machine Learning* berdasarkan konsep *Supervised Learning*. Algoritma KNN memberi asumsi kesamaan antara data baru dan kasus tersebut yang tersedia dan memasukkan kasus baru ke kategori yang sangat mirip dengan kategori yang ada atau tersedia. Algoritma KNN dapat menyimpan semua data yang tersedia dan mengklasifikasikan data baru berdasarkan atas kesamaan, yaitu ketika ada data baru muncul maka dapat dengan mudah diklasifikasikan ke dalam kategori bagus dengan tersebut.

Algoritma KNN juga dapat digunakan untuk konsep Regresi dan untuk Klasifikasi pola data tetapi sebagian besar kegunaannya digunakan untuk konsep masalah Klasifikasi. KNN merupakan jenis algoritma non-parametrik, yang maknanya tidak membangun asumsi apa saja pada data yang mendasarinya. Algoritma ini disebut algoritma dimana tidak langsung belajar dari dataset pelatihan yang ada melainkan menyimpan dataset yang ada dan pada saat klasifikasi, Algoritma ini melakukan tindakan mengekstraksidata membentuk pola pada dataset. Algoritma KNN pada fase membaca pola dataset pelatihan hanya menyimpan dataset tersebut dan ketika mendapatkan dataset baru, kemudian mengklasifikasikan dataset yang ada ke kategori yang mirip dengan dataset baru (Boukenze et al., 2012).

Secara umum, persamaan *Euclidean distance* pada 1-dimensional adalah sebagai berikut.

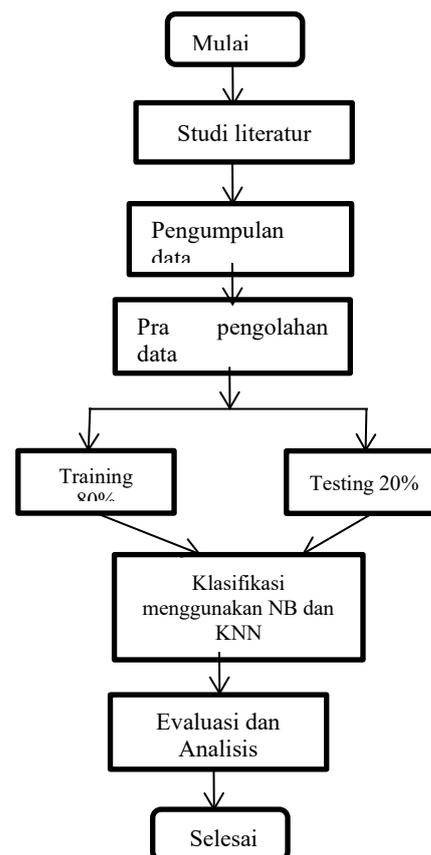
$$dis(x_1, x_2) = \sqrt{\sum_{i=0}^n (x_{1_i} - x_{2_i})^2}$$

Persamaan di atas dapat digunakan jika jumlah *independent variable* hanya ada satu variabel. jika ada banyak variabel yang digunakan mak menggunakan seperti di bawah ini.

$$dis = \sqrt{\sum_{i=0}^n (x_{1_i} - x_{2_i})^2 + (y_{1_i} - y_{2_i})^2 + \dots} \quad (3)$$

III. METODE PENELITIAN

Dalam penelitian ini menggunakan pendekatan metode kuantitatif dengan melalui tahapan yang ditunjukkan pada Gambar 1 dibawah ini:



Gambar 1 Tahapan Penelitian

Berdasarkan Gambar 1, tahapan penelitian dimulai dari studi literatur berupa membaca jurnal dan artikel, pengumpulan data yang diperoleh dari dataset publik. Sebelum dilakukan klasifikasi, terdapat proses pra

pengolahan data untuk mengecek apakah ada *missing value*, menghilangkan atribut yang tidak relevan.

Setelah itu dataset dibagi menjadi training dan testing dengan perbandingan 80% untuk training pembuatan model dan 20% untuk testing model yang telah dibuat.

Setelah dibagi dengan proporsi masing-masing, dataset akan diklasifikasikan menggunakan algoritma *Naive Bayes* dan K-NN untuk memprediksi label/class. Langkah terakhir yang dilakukan adalah evaluasi dan analisis kinerja model yang dibuat.

Tabel 1. Dataset penelitian

Atribut	Tipe	Deskripsi
A1-A10	Binary (0, 1)	Jawaban kode pertanyaan berdasarkan metode skrining yang digunakan
Age	Number	Usia balita(bulan)
Score by Q-chat-10	Number	1-10(kurang dari 3 tidak ASD. 3 atau lebih dari 3 ASD)
Sex	Character	Laki-laki atau perempuan
Ethnicity	String	Etnis
Born with jaundice	Boolean (yes or no)	Apakah memiliki penyakit kuning
Family member with ASD history	Boolean (yes or no)	Apakah ada anggota keluarga yang mengidap ASD
Who is completing the test	String	Jawaban diisi oleh orang tua, perawat, staf medis, dll
Class variable	String	ASD traits or No ASD traits (automatically assigned by the ASDTests app). (Yes / No)

Pada tabel 2 berisi 10 pertanyaan dari instrumen Q-CHAT-10 dengan masing-masing pertanyaan terdapat 4 opsi jawaban yang berbeda-beda. Berdasarkan keputusan dari Q-CHAT-10 yang dibuat, jika nilai yang diperoleh lebih dari 3 maka akan didiagnosis mengidap ASD.

Tabel 2 Detail pertanyaan pada Q-CHAT-10

Variabel	Pertanyaan
A1	Does your child look at you when you call his/her name?
A2	How easy is it for you to get eye contact with your child?
A3	Does your child point to indicate that s/he wants something? (e.g. a toy that is out of reach)
A4	Does your child point to share interest with you? (e.g. pointing at an interesting sight)
A5	Does your child pretend? (e.g. care for dolls, talk on a toy phone)
A6	Does your child follow where you're looking?

A. Sumber Data

Dataset yang digunakan pada penelitian ini merupakan dataset sekunder. Dataset ini diperoleh dari website kaggle (Thabtah, 2018). Jumlah atribut yang ada di dataset ini sebanyak 18 termasuk kelas dan jumlah data sebanyak 1054. Atribut dan deskripsinya ditunjukkan pada tabel 1. Pada tabel 1 terdapat atribut yang digunakan untuk mencatat identitas pasien dan 10 pertanyaan yang digunakan untuk mendeteksi ASD. Pertanyaan A1-A10 diperoleh dari instrumen Q-Chat 10 yang ditujukan untuk mendeteksi autisme pada balita yang berusia 18-24 tahun (NHS, 2012).

Variabel	Pertanyaan
A7	If you or someone else in the family is visibly upset, does your child show signs of wanting to comfort them? (e.g. stroking hair, hugging them)
A8	Would you describe your child's first words as:
A9	Does your child use simple gestures? (e.g. wave goodbye)
A10	Does your child stare at nothing with no apparent purpose?

B. Model

Model dibangun menggunakan Tools rapid Miner dimana tools ini dapat menangani keperluan data scientist. Sebelum model ini dibuat terdapat proses pra pengolahan data untuk mengecek apakah berisi missing value dan menghilangkan atribut-atribut yang tidak relevan. Hasil dari pra pengolahan data dijadikan sebagai pembuatan model dengan menerapkan algoritma Naive Bayes dan KNN.

C. Evaluasi dan Analisis

Model yang telah dibangun sebelumnya akan dievaluasi dari kinerjanya berdasarkan hasil akurasi, sensitivitas, dan spesifisitas.

IV. HASIL DAN PEMBAHASAN

Pada penelitian ini menggunakan dataset publik yang diperoleh dari Kaggle. Atribut yang dipakai dalam penelitian ini sebanyak 12. Salah satu diantaranya menjadi class atau label klasifikasi. Sebelas atribut lainnya yang dipilih berdasarkan ketentuan dari Q-CHAT-10 yaitu: A1-A10 dan 1 atribut yang lain adalah umur. Atribut-atribut yang dipilih pada penelitian ini disajikan pada Tabel 4.1.

Tabel 4.1 Atribut penelitian

No	Atribut	Value
1	A1	0,1
2	A2	0,1
3	A3	0,1
4	A4	0,1
5	A5	0,1
6	A6	0,1
7	A7	0,1
8	A8	0,1
9	A9	0,1
10	A10	0,1
11	Age (month)	12-36
12	Class	Yes, No

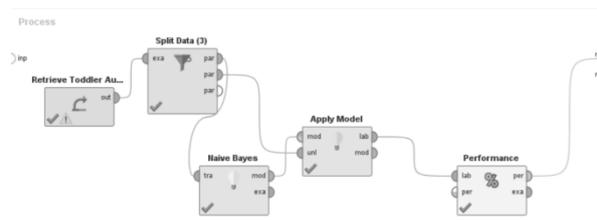
Sebelum masuk ke proses pembuatan model, terdapat proses pra pengolahan data yaitu mengecek apakah data ada *missing value* dan menghilangkan atribut-atribut yang tidak relevan. Jumlah keseluruhan atribut ada 19 termasuk class/label namun yang relevan sebanyak 12. Sepuluh atribut (A1-A10) dipilih berdasarkan pertanyaan yang ada di instrumen Q-CHAT-10, 1 atribut dipilih untuk mendeteksi ASD berdasarkan umur, dan 1 atribut menjadi kelas/label untuk hasil prediksi.

Jumlah data keseluruhan ini sebanyak 1054 dibagi menjadi data training dan testing. Pada penelitian ini pembagian data training dan testing masing-masing sebesar 80% training dan 20% testing. Dari pembagian tersebut diperoleh data training sebanyak 843

dan data testing sebanyak 211. Selanjutnya masuk ke proses pembuatan model Klasifikasi ASD. Pada penelitian ini menggunakan 2 algoritma klasifikasi yaitu Naive Bayes dan KNN guna membandingkan 2 kinerja yang dihasilkan oleh 2 algoritma tersebut.

1. Klasifikasi menggunakan Algoritma Naive Bayes

Pada Gambar 2 merupakan proses klasifikasi ASD menggunakan Algoritma *Naive Bayes*.



Gambar 2 Proses Klasifikasi ASD menggunakan Algoritma *Naive Bayes*

Setelah proses klasifikasi akan terbentuk model yang digunakan untuk mendeteksi ASD pada data testing. Model yang terbentuk dari Algoritma *Naive Bayes* dievaluasi dan dianalisis kinerjanya berdasarkan nilai akurasi, sensitifitas, dan spesifisitas.

accuracy: 94,79%

	true No	true Yes	class precision
pred. No	60	6	90.91%
pred. Yes	5	140	96.55%
class recall	92.31%	95.89%	

Gambar 3 Hasil Akurasi dari Naive Bayes

Pada gambar 3 menunjukkan bahwa nilai akurasi yang dihasilkan dari algoritma *Naive Bayes* sebesar 94,79%. Dari 211 data testing diantaranya 140 data terdeteksi benar positif ASD dan 60 data terdeteksi benar negatif ASD. Sedangkan 11 data lainnya misklasifikasi.

sensitivity: 95.89% (positive class: Yes)

	true No	true Yes	class precision
pred. No	60	6	90.91%
pred. Yes	5	140	96.55%
class recall	92.31%	95.89%	

Gambar 4 Hasil Sensitivitas dari *Naive Bayes*

Pada gambar 4 menunjukkan bahwa nilai sensitivitas yang dihasilkan dari algoritma *Naive Bayes* sebesar 95,89%. Dari 146 data testing yang berlabel positif ASD terklasifikasi benar positif ASD (TP) sebanyak 140. Sedangkan 6 data terklasifikasi salah (FN).

specificity: 92.31% (positive class: Yes)

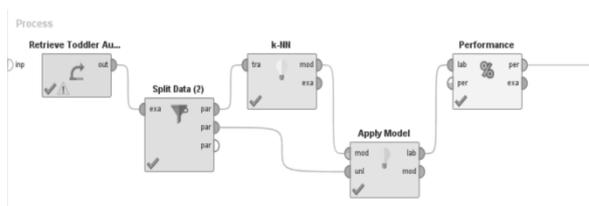
	true No	true Yes	class precision
pred. No	60	6	90.91%
pred. Yes	5	140	96.55%
class recall	92.31%	95.89%	

Gambar 5 Hasil Spesifisitas dari *Naive Bayes*

Pada gambar 4 menunjukkan bahwa nilai spesifisitas yang dihasilkan dari algoritma *Naive Bayes* sebesar 92,31%. Dari 65 data testing yang berlabel negatif ASD terklasifikasi benar negatif ASD (TN) sebanyak 60. Sedangkan 5 data terklasifikasi salah (FP).

2. Klasifikasi menggunakan Algoritma *K-Nearest Neighbor* (KNN)

Pada Gambar 6 merupakan proses klasifikasi ASD menggunakan KNN.



Gambar 6 Proses Klasifikasi ASD menggunakan Algoritma KNN

Setelah proses klasifikasi akan terbentuk model yang digunakan untuk mendeteksi ASD pada data testing. Selanjutnya model yang terbentuk dari algoritma KNN dievaluasi dan dianalisis kinerjanya berdasarkan nilai akurasi, sensitifitas, dan spesifisitas.

Tabel 3 Hasil Akurasi dari KNN berdasarkan nilai k

Akurasi	Nilai k		
	k = 5	k = 7	k = 9
	94,79%	95,73%	95,26%

Pada Tabel 3 menunjukkan bahwa KNN menghasilkan nilai akurasi yang berbeda tergantung nilai k yang dipilih. Klasifikasi menggunakan KNN dengan cara memasukkan beberapa nilai k yang dipilih. Pemilihan nilai k digunakan untuk mencari hasil akurasi yang baik untuk klasifikasi data testing tersebut. Dari ketiga nilai k tersebut yang menghasilkan nilai akurasi yang baik diperoleh pada nilai k = 7 dengan nilai akurasi sebesar 95,73%.

Tabel 4. Hasil Sensitivitas dari KNN berdasarkan nilai k

Sensitivitas	Nilai k		
	k = 5	k = 7	k = 9
	93,15%	93,84%	94,52%

Pada Tabel 4 menunjukkan bahwa KNN menghasilkan nilai sensitivitas yang berbeda. Dari ketiga nilai k tersebut, nilai akurasi yang baik diperoleh pada nilai k=9 sebesar 94,52%.

Tabel 5. Hasil Spesifisitas dari KNN berdasarkan nilai k

Spesifisitas	Nilai k		
	k = 5	k = 7	k = 9
	98,46%	100%	96,92%

Pada Tabel 5 menunjukkan bahwa KNN menghasilkan nilai spesifisitas yang berbeda. Dari ketiga nilai k tersebut, nilai akurasi yang baik diperoleh pada nilai k=7 sebesar 100%. Dengan hasil spesifisitas tersebut menunjukkan tidak ada misklasifikasi pada negatif ASD (TN). Untuk nilai kinerja yang optimal dari masing-masing pengukuran maka nilai k yang dipilih yaitu k = 7 dengan nilai akurasi sebesar 95,73%, sensitivitas sebesar 93,84%, dan spesifistas sebesar

100%. Selanjutnya keseluruhan kinerja kedua model dari Algoritma dan Naive Bayes dibandingkan untuk melihat kinerja yang terbaik dalam klasifikasi ASD pada data testing.

Tabel 6. Perbandingan kinerja dari *Naive Bayes* dan KNN

	NB	KNN
Akurasi	94,79%	95,73%
Sensitivitas	95,89%	93,84%
Spesifisitas	92,31%	100%

Berdasarkan tabel 6 hasil kinerja yang optimal diperoleh dari KNN dengan nilai $k = 7$. KNN menghasilkan nilai akurasi lebih baik dibandingkan NB dengan selisih 0,94%. Sedangkan nilai sensitivitas dari KNN lebih rendah daripada NB dengan selisih 2,05%. Untuk nilai spesifisitas dari KNN sangat bagus karena berhasil mengklasifikasikan data testing yang berlabel negatif ASD.

V. KESIMPULAN

Pada penelitian ini beberapa algoritma *machine learning* digunakan untuk deteksi dini ASD yang diterapkan pada dataset *Autism screening data for toddlers*. Algoritma yang digunakan yaitu: *Naive Bayes* dan KNN kemudian dibandingkan kinerja yang dihasilkan dari akurasi, sensitivitas, dan spesifisitas. Berdasarkan eksperimen yang telah dilakukan menunjukkan bahwa KNN memperoleh kinerja lebih tinggi daripada *Naive Bayes*. Hasil Akurasi, Sensitivitas, dan Spesifisitas yang diperoleh dari KNN sebesar 95,73%, 93,84%, dan 100%.

DAFTAR PUSTAKA

- A'yuniyah, Q., Tasia, E., Nazira, N., Pratama, P. F., Anugrah, M. R., Adhiva, J., & Mustakim, M. (2022). Implementasi Algoritma Naïve Bayes Classifier (NBC) untuk Klasifikasi Penyakit Ginjal Kronik. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 4(1), 72. <https://doi.org/10.30865/json.v4i1.4781>
- Boukenze, B., Mousannif, H., Haqiq, A., Sharma, D., Sharma, A., Mansotra, V., Rubiano, S. M. M., Garcia, J. A. D., Merchán, S. M., Duarte, J. A., Rubiano, S. M. M., García, J. D., Shahiri, A. M., Husain, W., Rashid, N. A., Marbán, Ó., Mariscal, G., Segovia, J., Saputra, R. A., ... Dahlan, E. A. (2012). Top 10 algorithms in data mining. *Knowledge and Information Systems*. <https://doi.org/10.1017/S0269888910000032>
- Katilik, A. N., & Djie, J. A. (2022). Penerapan Pendekatan Orff-Schulwerk untuk Meningkatkan Hasil Belajar Siswa dengan Autism Spectrum Disorder (ASD) dalam Pembelajaran Instrumen Ritmis Sederhana. *Seni Musik*, 12(1), 91–109.
- NHS. (2012). *Q-CHAT-10 Quantitative Checklist for Autism in Toddlers*. <https://www.autismalert.org/uploads/PDF/SCREENING--AUTISM--QCHAT-10 Question Autism Survey for Toddlers.pdf>
- Purwaningsih, E., & Nurelasari, E. (2021). Penerapan K-Nearest Neighbor Untuk Klasifikasi Tingkat Kelulusan Pada Siswa. *Syntax: Jurnal Informatika*, 10(01), 46–56. <https://doi.org/10.35706/syji.v10i01.5173>
- Riani, A., Susianto, Y., & Rahman, N. (2019). Implementasi Data Mining Untuk Memprediksi Penyakit Jantung Menggunakan Metode Naive Bayes. *Journal of Innovation Information Technology and Application (JINITA)*, 1(01), 25–34. <https://doi.org/10.35970/jinita.v1i01.64>
- Ridwan, A., Andono, P. N., & Supriyanto, C. (2018). Optimasi Klasifikasi Status Gizi Balita Berdasarkan Indeks Antropometri Menggunakan Algoritma Naive. *Teknologi Informasi*.
- Rishel, C. W., Greeno, C., Marcus, S. C., Shear, M. K., & Anderson, C. (2005). Use of the Child Behavior Checklist as a diagnostic screening tool in community mental health. *Research on Social Work Practice*, 15(3), 195–203.

- <https://doi.org/10.1177/1049731504270382>
- Sukamto, S., Adriyani, Y., & Aulia, R. (2020). Prediksi Kelompok UKT Mahasiswa Menggunakan Algoritma K-Nearest Neighbor. *JUITA: Jurnal Informatika*, 8(1), 121. <https://doi.org/10.30595/juita.v8i1.6267>
- Thabtah, F. A. (2018). *Autism screening data for toddlers*. <https://www.kaggle.com/datasets/fabdelja/autism-screening-for-toddlers>
- WHO. (2023). *Autism*. <https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders>
- Wu, X., Kumar, V., Ross, Q. J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z. H., Steinbach, M., Hand, D. J., & Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*. <https://doi.org/10.1007/s10115-007-0114-2>