

SYSTEM DESIGN OF AUTHICAL DISTURBANCE DIAGNOSIS IN CHILDREN USING THE K-NEAREST NEIGHBOR METHOD

Achya Puji Sari¹, Dedy Kurniadi², Sam Farisa Chaerul Haviana³

¹Jurusan Teknik Informatika

address : Jl. Raya Kaligawe KM.4 Semarang, email : achya@std.unissula.ac.id

²Jurusan Teknik Informatika

address : Jl. Raya Kaligawe KM.4 Semarang, email : ddy.kurniadi@unissula.ac.id

³Jurusan Teknik Informatika

address : Jl. Raya Kaligawe KM.4, email : ddy.kurniadi@unissula.ac.id

ABSTRACT

Autisme dapat dialami oleh anak dari berbagai ras, suku, strata sosial, dan ekonomi. Autisme merupakan gangguan perkembangan pervasif pada anak yang ditandai dengan adanya gangguan dan keterlambatan dalam bidang komunikasi, kognitif, perilaku, bahasa, dan interaksi sosial. Orang tua terkadang menganggap gangguan-gangguan tersebut sebagai keterlambatan perkembangan biasa namun pada kenyataannya jumlah penyandang spektrum autisme semakin meningkat. Menurut data dari badan kesehatan dunia (WHO) pada tahun 2009, prevalensi autisme di Indonesia mengalami peningkatan luar biasa, dari 1 per 1000 penduduk menjadi 8 per 1000 penduduk. Pada tahun 2009 dilaporkan bahwa jumlah anak penderita autisme mencapai 150-200 ribu. Salah satu cara agar orang tua dapat mengetahui anaknya adalah penderita autisme dengan menggunakan fasilitas pendeteksi. Penelitian ini dalam mendiagnosis autisme pada anak menggunakan metode K-Nearest Neighbor dengan menentukan parameter setting untuk nilai k. Di lakukan pengujian dengan black box testing dan confusion matrix, di dapat nilai akurasi tertinggi sebesar 95%, presisi 95.45%, recall 95.45%, f-measure 95.44%, pada nilai k=4.

Keywords: Autisme, K-Nearest Neighbor, confusion matrix,

This is an open access article under the CC BY license.



1. INTRODUCTION

Autisme dapat dialami oleh anak dari berbagai ras, suku, strata sosial, dan ekonomi. Autisme merupakan gangguan perkembangan pervasif pada anak yang ditandai dengan adanya gangguan dan keterlambatan dalam bidang komunikasi, kognitif, perilaku, bahasa, dan interaksi sosial. Orang tua terkadang kurang peka dengan gangguan-gangguan perkembangan tersebut. Mereka menganggap gangguan-gangguan itu sebagai keterlambatan perkembangan biasa. Namun, ketika perkembangan anak tidak menunjukkan kemajuan bahkan mengalami kemunduran orang tua baru sadar bahwa anaknya mengalami gangguan dalam perkembangan[1]. Jumlah penyandang spektrum autisme tampaknya semakin meningkat pesat. Beberapa negara terdapat kenaikan angka penderita autisme yang cukup tajam. Data Centre of Disease Control (CDC) di Amerika pada bulan Maret 2014, prevalensi (angka kejadian) Autisme adalah 1 dari 68 anak. Secara lebih spesifik 1 dari 42 anak laki-laki dan 1 dari 189 anak perempuan.. Sedangkan di Jepang dan Kanada sejak tahun 1980 jumlah kasus penderita autisme pertambahannya mencapai 40%. Sementara itu di California pada tahun 2002 terdapat 9 orang penderita autisme per-harinya, yang mana perkiraan prevalensi autisme adalah 4 sampai 5 anak per 10.000 anak[2]. Berdasarkan data yang dirilis Organisasi Kesehatan Dunia (WHO), prevalensi autisme di Indonesia

Receive : 10 Desember 2020; Revised From : 15 Desember 2020; Accepted : 27 Desember 2020;

Available Online : 1 Januari 2021

mengalami peningkatan luar biasa, dari 1 per 1000 penduduk menjadi 8 per 1000 penduduk dan melampaui rata-rata dunia yaitu 6 per 1000 penduduk. Pada tahun 2009 dilaporkan bahwa jumlah anak penderita autisme mencapai 150-200 ribu[3].

Saat ini hanya bisa dilakukan dengan datang ke psikolog, yang mana pastinya selain membutuhkan banyak biaya juga menjadikan waktu yang tidak *effesien* karena terkait dengan *procedural*, padahal pendeteksian awal itu penting, karena semakin dini gangguan *autism* pada anak terdiagnosa, maka semakin besar peluang untuk sembuh. Salah satu cara agar orang tua dapat mengetahui anaknya adalah penderita *autism* dengan menggunakan fasilitas pendeteksi.

Data mining dapat diaplikasikan dibidang kesehatan misalnya mendiagnosis penyakit kanker payudara, penyakit jantung, penyakit diabetes, gangguan autis, dan lain – lain[4]. Terdapat beberapa metode dalam mendiagnosis *autism* misalnya menggunakan K-Nearest Neighbor. Konsep mendasar dari algoritma KNN yaitu mencari jarak paling terdekat diantara data yang terevaluasi dengan sejumlah K tetangga (*neighbor*) paling dekat dalam data uji[4]. KNN bekerja dengan cara membandingkan data uji dan data *training/template*. Berdasarkan paparan sebelumnya maka dalam penelitian ini akan dibangun sebuah perancangan sistem diagnosa gangguan autis pada anak menggunakan metode *K-Nearest Neighbor* berbasis android. Pemilihan sistem operasi dikarenakan pertimbangan dari penelitian yang dilakukan oleh International Data Corporation (IDC) pada tahun 2015 menunjukkan bahwa android memiliki 82,8 % pangsa pasar di seluruh dunia, sementara iOS hanya memperoleh 13,9 %.

2. LITERATURE REVIEW

Dalam penyelesaian masalah yang mengenai cara mendeteksi autis secara dini, beberapa penelitian telah dilakukan guna menanggulangi dampak dari *autism*. Salah satunya adalah penelitian yang dilakukan oleh Gardenia dengan judul “Sistem Pakar Deteksi *Autisme* Pada Anak Menggunakan Metode *Fuzzy Tsukamoto*”. Penelitian ini memiliki latar belakang karena pertambahan penderita autis setiap tahunnya, dirasa perlu untuk meningkatkan pengetahuan dan mendeteksi penderita autis dengan cepat untuk melakukan prevensi. Dengan menggunakan algoritma *Fuzzy Tsukamoto*, sistem ini dapat mendeteksi penyakit autis dengan akurasi sampai ke angka 73,33% dari 40 data sampel dan data yang diuji sebanyak 15 data dengan 11 hasil diagnosis yang tepat. Hal ini dapat terjadi karena perbedaan antara permasalahan yang nyata terlihat oleh pakar dan juga pengalaman yang pakar telah alami, hingga beberapa kondisi yang belum disediakan dari sistem dari aturan yang digunakan [5]

Pada penelitian yang dilakukan oleh Budiman pada pendeteksi Jenis Autis pada Anak Usia Dini Menggunakan Metode *Linear Discriminant Analysis* (LDA). Pada penelitian tersebut membagi autis mejadi tiga jenis yaitu kelas ringan, sedang, dan berat sesuai dari keadaan penderita *autisme*. Dalam melakukan deteksi dari *autisme* ini dapat menggunakan metode *linear discriminant analysis* (LDA) untuk mendapatkan hasil akurasi yang baik. Dengan menggunakan 75 data latih, sistem ini dapat menghasilkan nilai akurasi sebesar 88%. [6]

Penelitian sebelumnya meneliti tentang Prediksi Ketepatan Waktu Lulus Mahasiswa dengan *K-Nearest Neighbor* dan *Naïve Bayes Classifier* (Studi Kasus Prodi D3 Sistem Informasi Universitas Airlangga). Penelitian ini menawarkan penggunaan metode penggalian data untuk memprediksi waktu lulus mahasiswa menggunakan dua metode yaitu *K-Nearest Neighbor* dan *Naïve Bayes Classifier*. Hasil dari penelitian ini berupa sistem yang dapat memprediksi ketepatan waktu lulus. Uji coba dilakukan dengan menggunakan data lulusan mahasiswa D3 Sistem Informasi Universitas Airlangga. Hasil uji coba menunjukkan bahwa metode *K-Nearest Neighbor* menghasilkan akurasi lebih tinggi dibandingkan dengan *Naïve Bayes Classifier*. Akurasi tertinggi diperoleh dengan menggunakan metode *K-Nearest Neighbor* yaitu sebesar 98.7%. Oleh karena itu dapat disimpulkan bahwa sistem yang dibangun pada penelitian ini mampu memprediksi ketepatan waktu lulus dengan akurasi cukup tinggi [7]

Pada penelitian yang berjudul Sistem Pakar Diagnosa Penyakit Demam : DBD, Malaria, dan Tifoid Menggunakan Metode *K-Nearest Neighbor – Certainty Factor*. Metode yang digunakan untuk mendukung sistem pakar tersebut adalah metode *K-Nearest Neighbor – Certainty Factor* yang merupakan penggabungan 2 metode dimana hasil klasifikasi dari metode *K Nearest Neighbor* akan di beri nilai kepastian oleh metode *Certainty Factor* sehingga menghasilkan suatu diagnosis penyakit. Pada penelitian ini data latih dan data uji yang di gunakan berjumlah 143 data. Berdasarkan hasil pengujian variasi nilai K di dapatkan akurasi sebesar 88.37%. Pada pengujian variasi data latih di dapatkan akurasi sebesar 86.04%. Pada pengujian rasio data latih dan data uji di dapatkan akurasi sebesar 95%. Pada pengujian variasi jumlah data uji didapatkan akurasi sebesar 90%. Pada pengujian variasi data uji di dapatkan rata-rata akurasi sebesar 97.22%. Pada pengujian perbandingan metode, metode *k-nearest neighbor-certainty factor* menghasilkan akurasi sebesar 84.79% [8]

3. RESEARCH METHOD

3.1. DATA UNDERSTANDING

Pada tahapan data understanding akan dilakukan pengumpulan data, kemudian mempelajari data tersebut dengan tujuan untuk mengenal data, dan melakukan identifikasi dari data. Data yang digunakan merupakan data yang terdapat pada uci machine learning repository yang mana terdiri dari 280 instances dan 17 feature. Data akan di bagi menjadi data training dan data testing. Data set yang disediakan pada uci machine learning dalam format ARFF, dimana format ini merupakan format yang dapat dibaca oleh tools WEKA.

3.2. DATA PREPARATION

Dalam melakukan data preparation, akan dilakukan beberapa tahapan yaitu data cleaning dan data reduction.

- 1) *Data cleaning* dapat digunakan untuk data yang *missing value*. Karena ditemukan adanya data yang terlewat tidak terisi (*missing value*) pada data. Cara pengisian atribut yang kosong adalah dengan memberikan nilai berdasarkan nilai yang paling banyak atau dominan dalam atribut tersebut.
- 2) *Data reduction* digunakan untuk menghasilkan data set yang volumenya lebih kecil. Salah satu strategi *data reduction* yang digunakan pada penelitian ini adalah dengan mereduksi ukuran data set atau menghilangkan atribut-atribut yang tidak relevan / berpengaruh terhadap hasil dari algoritma K-Nearest Neighbor.

3.3. MODELING

Model yang akan digunakan dalam tahap ini menggunakan algoritma K-Nearest Neighbor dengan menghitung jarak euclidean dan menentukan parameter setting.

$$d_i = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2} \quad (1)$$

Keterangan :

d_i = jarak kedekatan

p = jumlah seluruh atribut data

x_1 = data latih (data training)

x_2 = data uji (data testing)

i = variabel data ke- i

1. Setelah itu mengurutkan berdasarkan nilai *Euclidean distance*.
2. Menentukan *k record* klasifikasi terdekat.
3. Target *output* merupakan kelas yang mayoritas.

3.4. EVALUATION

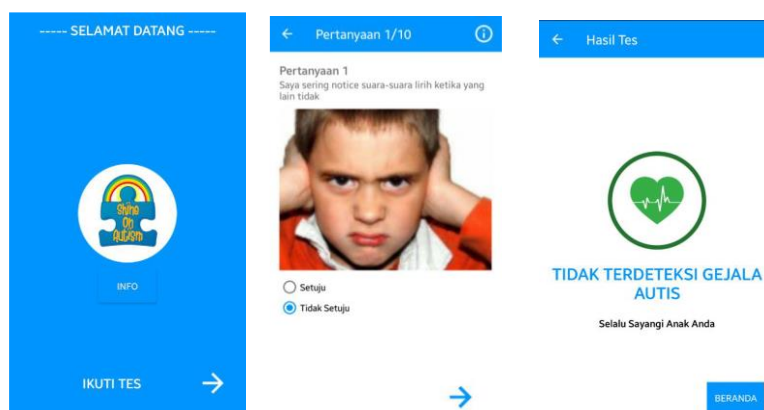
Pada penelitian ini menggunakan confusion matrix sebagai pengujian akurasi dari algoritma yang diterapkan.

Tabel 3.1 Confusion matrix

		Nilai Sebenarnya	
		True	False
Nilai Prediksi	True	TP (True Positive) Correct Result	FP (False Positive) Unexpected Result
	False	FN (False Negative) Missing Result	TN (True Negative) Corect Absence of Result

4. RESULT AND ANALYSIST

Berikut merupakan hasil dari implementasi sistem diagnosa gangguan autisme pada anak menggunakan metode *k-nearest neighbor* :



Gambar 4.1 User Interface Aplikasi

Tabel 4.1 Perhitungan Confusion Matrix

Nilai k	Akurasi	Presisi	Recall	F-measure
2	93.75%	95.45%	93.33%	94.37%
3	92.5%	90.9%	95.23%	93.01%
4	95%	95.45%	95.45%	95.44%
5	93.75%	90.9%	97.56%	94.11%
7	91.25%	88.56%	95.12%	91.76%

5. CONCLUSION

Setelah melakukan penelitian mengenai deteksi gangguan autis pada anak menggunakan algoritma K-Nearest Neighbor, penulis dapat menarik kesimpulan bahwa alorigma K-Nearest Neighbor yang diterapkan dalam penelitian ini dapat mendeteksi penyakit autis dengan akurasi sampai ke angka 95% pada nilai k = 4 dari 200 data sampel dan data yang diuji sebanyak 80 data dengan 76 hasil diagnosis yang tepat. Presisi yang di dapat 95,45%, recall 95,45%, F-measure 95,44%. Sedangkan pada sistem dilakukan juga pengujian menggunakan black box testing yang mana pada pengujian ini sistem dapat mencapai hasil output sesuai dengan yang diharapkan

REFERENCES

- [1] Danuadmaja, Bonny. 2003. "Terapi Anak Autis di Rumah". Jakarta : Puspa Swara. Cetakan 1. Hal. 2-6.
- [2] Safaria, Triantoro. 2005. "Autisme : Pemahaman Baru Untuk Hidup Bermakna Bagi Orang Tua". Yogyakarta : Graha Ilmu. Cetakan 1.
- [3] Yuwono, Joko. 2009. "Memahami Dunia Anak Autistik". Bandung : Alfabeta.
- [4] D. T. Larose. 2005. "Discovering Knowledge in Data : An Introduction to Data Mining". John Willey dan Sons Inc.
- [5] M. Gardenia, T. Tursina, and H. Sastypratiwi. 2016. "Sistem Pakar Deteksi Autisme Pada Anak Menggunakan Metode Fuzzy Tsukamoto". *J. Sist dan Teknol Inf.* vol 4. no 1. pp 33–38.
- [6] E. Budiman, E. Santoso, and T. Afrianto. 2017. "Pendeteksi Jenis Autis pada Anak Usia Dini Menggunakan Metode Linear Discriminant Analysis (LDA)". vol 1. no 7. pp. 583–592.
- [7] W. I. Sabilla and T. E. Putri. 2017. "Prediksi Ketepatan Waktu Lulus Mahasiswa dengan K-Nearest Neighbor dan Naïve Bayes Classifier". *J Komp Terap.* vol 3. no 2. pp 233–240.
- [8] C. S. Fatoni and F. D. Noviantha. 2018. "Case Based Reasoning Diagnosis Penyakit Diferi dengan Algoritma K-Nearest Neighbor". *Creat Inf Technol J.* vol 4. no 3. pp 220–232.