

PENERAPAN *SUPPORT VECTOR MACHINE (SVM)* PADA *SMALL DATASET* UNTUK DETEKSI DINI GANGGUAN AUTISME

Bayu Sugara¹; Agus Subekti²

^{1,2}Program Studi Magister Ilmu Komputer
STMIK Nusa Mandiri Jakarta
<http://nusamandiri.ac.id/>

¹bayusugaraa@gmail.com, ²agus@nusamandiri.ac.id

²Pusat Penelitian Elektronika dan Telekomunikasi
Lembaga Ilmu Pengetahuan Indonesia (LIPI)
<http://lipi.go.id/>
agus@subekti@lipi.go.id

Abstract— *As time goes by the development of science and information technology, the presence of machine learning in the field of computers have become one of the trends and attracts a lot of attention. The application of machine learning is not regardless of using the data. The large data is often used in the learning process of machine learning. The development of machine learning grows so rapid that can enable the large data be accumulated so quickly. However, the machine learning is rarely found using the small data in the learning process. These small datasets are usually private that taken from the organization which would be the object of research such as the data on the banks, the hospitals, the factories and the service companies. In this research, the researcher exploit the Support Vector Machine of Algorithm and K-Fold Corss Validation to test the value of the accuracy of the small datasets and to utilize the ensemble techniques to determine how the ensemble technique affected the Support Vector Machine of Algorithm. The result of this research shows that the ensemble technique can improve the accuracy performance in Support Vector Machine. The SVM Algorithm model and the ensemble technique with poly kernel shows the best accuracy value of 91 %.*

Keywords: *Machine Learning, Support Vector Machine, Small Dataset, Poly kernel, Teknik Ensemble.*

Intisari— Seiring dengan perkembangan ilmu pengetahuan dan teknologi informasi, kehadiran *machine learning* dibidang komputer telah menjadi salah satu tren dan menarik banyak perhatian. Penggunaan *machine learning* tidak terlepas dari penggunaan data dalam pembelajarannya. Data yang besar merupakan data yang sering digunakan dalam proses pembelajaran *machine learning*. Perkembangan

machine learning yang sangat pesat dapat memungkinkan data yang besar cepat pula terakumulasi. Namun, jarang ditemukan *machine learning* menggunakan data yang kecil (*small dataset*) dalam proses pembelajarannya. *Small dataset* ini biasanya bersifat *private* yang diambil dari sebuah organisasi yang akan dijadikan objek penelitian seperti data bank, rumah sakit, pabrik dan perusahaan jasa. Dalam penelitian ini peneliti menggunakan algoritma *Support Vector Machine* dan *k-fold corss validation* untuk menguji nilai keakuratan *small dataset* serta menggunakan teknik *ensemble* untuk mengetahui seberapa pengaruhnya teknik *ensemble* terhadap algoritma *Support Vector Machine*. Hasil dari penelitian ini menunjukkan bahwa teknik *ensemble* dapat meningkatkan performa akurasi pada *Support Vector Machine*. Model algoritma *Support Vector Machine* dan teknik *ensemble* dengan *poly kernel* menunjukkan nilai akurasi terbaik yaitu sebesar 91%.

Kata Kunci: *Machine Learning, Support Vector Machine, Small Dataset, Poly kernel, Teknik Ensemble.*

PENDAHULUAN

Tumbuh kembang anak merupakan hal yang sangat penting bagi orang tua (Permono, 2013). Dengan melakukan deteksi dini terhadap tumbuh kembang anak, orang tua diharapkan mampu mengetahui tumbuh kembang pada anaknya baik tumbuh kembang yang normal maupun adanya gangguan dalam tumbuh kembang anak. Salah satu gangguan tumbuh kembang yang dialami oleh anak adalah gangguan autis.

Autis merupakan gangguan perkembangan fungsi otak yang mencakup bidang sosial, komunikasi verbal (bahasa) dan non-verbal,

imajinasi dan fleksibilitas, lingkup minat, kognisi dan perhatian (Gardenia, Tursina, & Pratiwi, 2015). Anak autistik ditinjau dari masa kemunculannya atau kejadiannya dapat terjadi sejak lahir yang disebut dengan autistik klasik dan sesudah lahir dimana anak hingga umur 1-2 tahun menunjukkan perkembangannya yang normal (Siwi & Anganti, 2017).

Autisme memiliki dampak yang nyata di sisi sosial dari penderita. Penderita penyakit ini seakan-akan hidup diduniannya sendiri (Hendita, Kusuma, Si, Cs, & Oktana, 2012). Dengan adanya hal ini, penyandang autis akan terisolir dengan dunia luar sehingga banyak orang tua yang malu dan kurang percaya diri dengan keadaan anaknya yang memiliki kekurangan sehingga banyak orang tua juga tidak bisa menerima keadaan yang dimiliki anaknya (Susanto, 2014). Penyandang autis juga memiliki keterbatasan dalam hal komunikasi sehingga menyebabkan mereka berbicara tidak dipakai untuk berkomunikasi dengan orang lain tetapi dengan dirinya sendiri (Boham, 2013).

Hal ini lah yang dianggap perlu untuk ditangani, terhubung 1 dari 600 anak di Indonesia menyandang penyakit autis, harus ada suatu cara untuk menyelesaikan permasalahan (Budiman, Santoso, & Afirianto, 2017). Ini akan sangat bermanfaat bagi dokter, membantu mereka mendeteksi *Autism Spectrum Disorder* pada tahap yang jauh lebih produktif (Bekerom, 2012). Semakin cepat dokter dalam pengambilan keputusan, semakin cepat pula anak terkena autis dapat ditangani.

Dengan minimnya informasi autisme yang belum didapatkan oleh masyarakat luas khususnya orang tua, banyak yang tidak mengerti apa itu gangguan autisme (Merianto, 2016) dan bagaimana cara penanganannya. Para orang tua lebih memilih mencari seorang dokter/ahli tumbuh kembang anak untuk berkonsultasi (Sugara, Widyatmoko, Prakoso, 2018). Dengan ketersediaan dokter/ahli tumbuh kembang anak saat ini, serta jika menggunakan prosedur pendiagnosa secara lengkap membuat orang tua mengeluarkan biaya yang relatif besar serta waktu yang terbuang hanya untuk berkonsultasi.

Seiring dengan perkembangan ilmu pengetahuan dan teknologi informasi, kehadiran *machine learning* di bidang komputer telah menarik banyak perhatian. *Machine learning* menjadi sebuah tren dalam teknologi informasi. Saat ini perkembangan teknologi informasi telah merambah ke berbagai sektor termasuk kesehatan (Munawarah, Soesanto, & Faisal, 2016). *Machine learning* yang digunakan pada bidang kesehatan bertujuan untuk memprediksi sebuah penyakit. Dengan hasil prediksi yang lebih cepat diterima

oleh dokter/pakar kesehatan, semakin cepat pula tindakan atau penangannya terhadap sebuah penyakit.

Penggunaan *machine learning* tidak lepas dari penggunaan data itu sendiri. *Machine learning* memainkan peran yang luas dalam pengembangan terutama dalam pengembangan data analitik (Alarifi & Young, 2018). Data yang besar merupakan data yang sering digunakan dalam proses pembelajaran *machine learning* itu sendiri. Perkembangan *machine learning* yang sangat pesat dapat memungkinkan data yang besar cepat pula terakumulasi. Namun, jarang ditemukan *machine learning* menggunakan data yang kecil (*small dataset*) dalam proses pembelajarannya. *Small dataset* ini biasanya bersifat *private* yang diambil dari sebuah organisasi yang akan dijadikan objek penelitian seperti data bank, rumah sakit, pabrik dan perusahaan jasa.

Penelitian dengan *small dataset* pernah dilakukan oleh (Shaikhina et al., 2015), *Dataset* eksperimental dalam bioteknologi umumnya terbatas dalam ukuran, sehingga membuat *Machine Learning* (ML) tidak praktis untuk pemodelan prediktif. Teknik-teknik baru dari *multiple run* untuk pengembangan model dan analisis data pengganti untuk validasi model disarankan untuk prediksi hasil biomedis berdasarkan pada *dataset* kecil untuk tugas klasifikasi dan regresi. Kerangka yang diusulkan diaplikasikan untuk merancang model *Neural Network* untuk stratifikasi risiko fraktur tulang *osteoarthritic*, dan model *Decision Tree* untuk prediksi penolakan transplantasi ginjal yang dimediasi oleh antibodi. Meskipun *set data* kecil (35 spesimen tulang dan 80 transplantasi ginjal), kedua model mencapai akurasi tinggi masing-masing 98,3% dan 85%.

Penelitian selanjutnya dengan menggunakan *small dataset* telah dilakukan oleh (Sharma & Sharma, 2016), Klasifikasi adalah teknik pembelajaran yang dilindungi yang penting yang digunakan oleh banyak aplikasi. Faktor penting pada dimana kinerja dari sebuah *classifier* tergantung adalah ukuran *dataset* yang akan digunakan oleh *classifier* untuk dilatih. Di naskah ini penulis telah menganalisis lima berbeda teknik klasifikasi (yaitu pohon keputusan, KNN, SVM, linear diskriminan dan metode *Ensemble*) dalam hal AUC dan akurasi prediksi ketika dilatih menggunakan *dataset* kecil dengan dimensi yang berbeda. Penelitian telah dilakukan menggunakan *dataset* dengan 24 fitur dan 400 *instance* (sampel). Hasilnya menunjukkan bahwa, secara umum metode *ensemble* (menggunakan meningkatkan pohon) berkinerja lebih baik daripada yang lain tetapi kinerja sedikit terdegradasi dengan dimensi berkurang.

Sedangkan penelitian yang dilakukan oleh (Feng, Zhou, & Dong, 2019) menggunakan *small dataset* untuk memprediksi cacat pematatan oleh regresi DNN dengan *dataset* kecil yang berisi 487 titik data. Itu ditemukan bahwa DNN pra-terlatih dan *fine-tuned* menunjukkan kinerja generalisasi yang lebih baik melalui jaringan saraf dangkal, mendukung mesin vektor, dan DNN dilatih oleh metode konvensional. DNN terlatih mengubah eksperimen yang tersebar data menunjuk ke peta akurasi tinggi dalam kimia dimensi tinggi dan parameter pemrosesan ruang. Meskipun DNN dengan *dataset* besar adalah solusi optimal, DNN dengan *dataset* kecil dan pra-pelatihan bisa pilihan yang masuk akal ketika *dataset* besar tidak tersedia dalam studi materi.

Berdasarkan uraian tersebut diatas, penelitian deteksi dini gangguan autisme ini akan menggunakan *small dataset* dengan salah satu algoritma *machine learning* yaitu dengan algoritma *support vector machine* yang bertujuan untuk memperoleh nilai akurat melalui permodelan *machine learning* dalam deteksi dini gangguan autisme. Penelitian deteksi dini gangguan autisme ini juga menggunakan teknik *ensemble* yang akan diuji terhadap algoritma *support vector machine* untuk dapat menjawab pertanyaan sebagai berikut:

1. Bagaimana efektifitas penerapan algoritma *support vector machine* dengan menggunakan *small dataset*?
2. Bagaimana nilai akurasi yang dihasilkan dari penerapan algoritma *support vector machine* dengan menggunakan *small dataset*?
3. Bagaimana pengaruh teknik *ensemble* pada pengujian algoritma *support vector machine* dengan menggunakan *small dataset*?

BAHAN DAN METODE

A. Algoritma Support Vector Machine

Pengertian *Support Vector Machine* (SVM) yaitu sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi – fungsi linier dalam sebuah fitur yang berdimensi tinggi dan dilatih dengan menggunakan algoritma pembelajaran yang didasarkan pada teori optimasi (Puspitasari, Ratnawati, & Widodo, 2018). *Support Vector Machine* (SVM) diperkenalkan oleh Vapnik, Boser dan Guyon pada tahun 1992 (Fridayanthie, 2015) sebagai rangkaian dari beberapa konsep-konsep unggulan dalam bidang *pattern recognition* (Susilowati, Sabariah, & Gozali, 2015).

Tingkat akurasi pada model yang akan dihasilkan oleh proses peralihan dengan SVM sangat bergantung terhadap fungsi kernel dan parameter yang digunakan (Parapat & Furqon, 2018). Berdasarkan dari karakteristiknya, metode

SVM dibagi menjadi dua, yaitu SVM Linier dan SVM Non-Linier. SVM linier merupakan data yang dipisahkan secara linier, yaitu memisahkan kedua class pada hyperplane dengan soft margin. Sedangkan SVM Non-Linier yaitu menerapkan fungsi dari kernel trick terhadap ruang berdimensi tinggi (Rachman & Purnami, 2012).

Pada dokumentasi Matlab, SVM untuk prediksi disebut dengan SVM Regression terdiri dari fungsi linear dan nonlinear dengan primal formula dan dual formula) (Nurmasani, Utami, & Al Fatta, 2017). Metode SVM Regression disebut sebagai suatu teknik nonparametric karena bergantung pada fungsi kernel. Beberapa tipe kernel yang digunakan dapat dijabarkan pada tabel 1 berikut:

Tabel 1. Tipe Kernel

Nama Kernel	Fungsi Kernel
Linear (dot)	$G(x_1, x_2) = x_1'x_2$
RBF	$G(x_1, x_2) = exp$
Polynomial	$G(x_1, x_2) =$

Sumber: (Nurmasani et al., 2017)

Matriks Gram adalah matriks n x n yang mengandung unsur $g_{i,j} = G(x, z)$. Setiap elemen $g_{i,j}$ sama dengan nilai awal dari suatu prediktor dengan nilai peubah φ . Namun, tidak perlu mengetahui nilai φ , karena dapat menggunakan fungsi kernel untuk menghasilkan matriks Gram secara langsung. Dengan menggunakan metode ini, SVM nonlinier menemukan fungsi optimal $f(x)$ dalam perubahan ruang prediktor.

Rumus ganda untuk SVM regresi nonlinear menggantikan nilai awal prediktor $X^T Z$ dengan elemen yang sesuai dari matriks Gram ($g_{i,j}$). SVM regresi nonlinier menemukan koefisien yang meminimalkan seperti pada persamaan 1:

$$L(a) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i - \alpha_j) (\alpha_j - \alpha_i) G(x_i, x_j) + \varepsilon \sum_{i=1}^N (\alpha_i + \alpha_i) - \sum_{i=1}^N y_i (\alpha_i + \alpha_i) \dots \dots \dots (1)$$

Dimana:

$$\sum_{n=1}^N (\alpha_n - \alpha_n) = 0 \dots \dots \dots (2)$$

$$\forall n: 0 \leq \alpha_n \leq C \dots \dots \dots (3)$$

$$\forall n: 0 \leq \alpha_n \leq C(4) \dots \dots \dots (4)$$

Kondisi Karush-Kuhn-Tucker (KKT) harus memenuhi kondisi sebagai berikut:

$$\forall n: \alpha_n (\varepsilon + \xi_n - y_n + f(x_n)) = 0 \dots \dots \dots (5)$$

$$\forall n: \alpha_n (\varepsilon + \xi_n + y_n - f(x_n)) = 0 \dots \dots \dots (6)$$

$$\forall n: \xi_n(C - \alpha_n) = 0 \dots\dots\dots (7)$$

$$\forall n: \xi_n (C - \alpha_n) = 0 \dots\dots\dots (8)$$

Kondisi tersebut menunjukkan bahwa epsilon memiliki fungsi pengganda Lagrange $\alpha_n = 0$ dan $\alpha_n = 0$. Pengganda Lagrange bukan nol disebut dukungan *vector* atau *support vector*. Fungsi yang digunakan untuk memprediksi nilai baru pada *support vector* seperti persamaan 9 berikut:

$$f(x) = \sum_{n=1}^N (\alpha_n - \alpha_n) G(x_n, x) + b \dots\dots\dots (9)$$

B. Teknik Pengumpulan Data

Data yang digunakan dalam penelitian deteksi dini gangguan autisme ini merupakan data primer yang diperoleh dari hasil wawancara disebuah lembaga Autisme di Bekasi Jawa Barat, selain wawancara, data penelitian ini juga diperoleh dari hasil sebaran kuesioner online yang dapat diakses dengan menggunakan url <https://forms.gle/YSvYcnRxi21dLzfP9>.

Data yang diperoleh terdiri dari 2 parameter yaitu data gejala autisme yang digunakan sebagai atribut pada penelitian ini sedangkan data gangguan autisme digunakan sebagai *class* yang dijelaskan sebagai berikut:

Tabel 2. Data Gangguan Autisme

Kode Gangguan	Nama Gangguan
GG01	Gangguan Perilaku
GG02	Gangguan Komunikasi
GG03	Gangguan Interaksi Sosial

Sumber: (Sugara & Subekti, 2019)

Sedangkan yang data primer yang dijadikan sebagai atribut pada penelitian deteksi dini gangguan autisme ini dapat dijelaskan sebagai berikut:

Tabel 3. Data Gejala Autisme

Kode Gejala	Nama Gejala
GJ01	Tidak memiliki kontak mata
GJ02	Suka diam/menyendiri
GJ03	Tidak suka dipeluk
GJ04	Tidak dapat merespon jika dipanggil orang
GJ05	Suka melakukan kegiatan/gerakan secara berulang-ulang
GJ06	Suka terpaku terhadap benda-benda tertentu
GJ07	Suka menyukai hal yang aneh seperti mencium-cium benda
GJ08	Suka mengungkapkan emosi (sedih, senang, marah dll) dengan sendirinya tanpa sebab
GJ09	Tidak bisa diam
GJ10	Tidak dapat berbicara
GJ11	Bisa berbicara namun tidak jelas
GJ12	Sering berbicara berlebihan
GJ13	Suka mengucapkan bahasa/kata-kata yang aneh secara berulang-ulang
GJ14	Tidak dapat menunjuk sesuatu dengan jari

Kode Gejala	Nama Gejala
	sendiri
GJ15	Tidak dapat menunjukkan keinginan dengan kata-kata
GJ16	Suka menarik-narik orang lain jika menginginkan sesuatu
GJ17	Tidak ada usaha dalam berkomunikasi
GJ18	Menghindar jika didekati
GJ19	Tidak dapat berinteraksi dengan lingkungan sekitar
GJ20	Tidak tertarik dengan orang lain
GJ21	Tidak peduli dengan sekitarnya
GJ22	Tidak suka dengan keramaian
GJ23	Tidak suka bermain dengan teman sebayanya
GJ24	Tidak dapat bersosialisasi dengan orang lain

Sumber: (Sugara & Subekti, 2019)

HASIL DAN PEMBAHASAN

Pengujian deteksi dini gangguan autisme ini menggunakan 67 dataset yang diperoleh dari hasil wawancara dan sebaran kuesioner. Data set yang akan diuji pada penelitian ini adalah sebagai berikut:

Tabel 4. Dataset Gangguan Autisme

	Anak Ke-1	Anak Ke-2	Anak Ke-3	Anak Ke-4	Anak Ke-5
GJ01	Tidak	Ya	Tidak	Tidak	Tidak
GJ02	Ya	Tidak	Ya	Ya	Ya
GJ03	Tidak	Tidak	Tidak	Ya	Ya
GJ04	Tidak	Ya	Tidak	Tidak	Tidak
GJ05	Ya	Ya	Tidak	Ya	Ya
GJ06	Ya	Tidak	Ya	Ya	Tidak
GJ07	Ya	Tidak	Ya	Tidak	Ya
GJ08	Tidak	Tidak	Ya	Ya	Ya
GJ09	Tidak	Ya	Ya	Tidak	Tidak
GJ10	Tidak	Ya	Tidak	Tidak	Tidak
GJ11	Ya	Tidak	Ya	Tidak	Ya
GJ12	Ya	Tidak	Tidak	Ya	Ya
GJ13	Ya	Tidak	Ya	Ya	Tidak
GJ14	Tidak	Ya	Tidak	Tidak	Ya
GJ15	Tidak	Ya	Ya	Tidak	Tidak
GJ16	Ya	Tidak	Tidak	Tidak	Tidak
GJ17	Tidak	Ya	Ya	Tidak	Ya
GJ18	Tidak	Tidak	Tidak	Tidak	Tidak
GJ19	Ya	Ya	Ya	Ya	Ya
GJ20	Ya	Ya	Ya	Ya	Tidak
GJ21	Ya	Ya	Ya	Ya	Ya
GJ22	Ya	Tidak	Tidak	Ya	Ya
GJ23	Ya	Ya	Tidak	Ya	Ya
GJ24	Ya	Ya	Ya	Ya	Ya
Hasil	GG03	GG02	GG01	GG03	GG03

Sumber: (Sugara & Subekti, 2019)

Penelitian deteksi dini gangguan autisme ini menggunakan model algoritma *Support Vector Machine* (SVM) dengan menggunakan pengujian *k-fold cross validation*. Dalam penelitian ini digunakan beberapa kernel SVM yang digunakan untuk menentukan akurasi terbaik diantaranya adalah *normalized poly kernel*, *poly kernel* dan *RBF kernel*. Diuji kembali dengan menggunakan teknik *ensemble* yang terdiri dari AdaBoost dan Bagging

untuk menentukan nilai akurasi terbaik. Hasil dari pengujian kernel SVM dapat dijelaskan sebagai berikut:

A. Hasil Pengujian Algoritma SVM

Dari pengujian algoritma SVM dengan menggunakan beberapa teknik kernel dapat dihasilkan sebagai berikut:

Tabel 5. Pengujian SVM

	Poly Kernel	Normalized Poly Kernel	RBF Kernel
Akurasi	83.58	85.07	56.71
TP Rate	0.836	0.851	0.567
FP Rate	0.123	0.165	0.567
Precision	0.832	0.854	0
Recall	0.836	0.851	0.567
F-Measure	0.834	0.829	0
MCC	0.717	0.745	0
ROC Area	0.894	0.861	0.501
PRC Area	0.796	0.77	0.417

Sumber: (Sugara & Subekti, 2019)

Dari Tabel 5. Pengujian SVM diatas dapat dijelaskan bahwa *Normalized Poly Kernel* memiliki nilai akurasi terbaik sebesar 85.07% dengan TP Rate sebesar 0.851 dan FP Rate sebesar 0.165.

B. Hasil Pengujian Algoritma SVM + Teknik Ensemble AdaBoost

Selanjutnya pengujian dengan menggunakan SVM dan salah satu teknik *ensemble* yaitu AdaBoost. Hasil pengujiannya sebagai berikut:

Tabel 6. Pengujian SVM + Teknik Ensemble Adaboost

	Poly Kernel	Normalized Poly Kernel	RBF Kernel
Akurasi	85.07	79.1	56.71
TP Rate	0.851	0.791	0.567
FP Rate	0.119	0.151	0.567
Precision	0.849	0.782	0
Recall	0.851	0.791	0.567
F-Measure	0.849	0.785	0
MCC	0.744	0.641	0
ROC Area	0.912	0.939	0.501
PRC Area	0.853	0.903	0.417

Sumber: (Sugara & Subekti, 2019)

Dari Tabel 6. Pengujian SVM dengan salah satu teknik *ensemble* yaitu AdaBoost diatas dapat dijelaskan bahwa *Poly Kernel* memiliki nilai akurasi terbaik sebesar 85.07% dengan TP Rate sebesar 0.851 dan FP Rate sebesar 0.119. Teknik *Esemble* AdaBoost ini menunjukkan dapat meningkatkan nilai akurasi terhadap perhitungan SVM menggunakan *poly kernel*.

C. Hasil Pengujian Algoritma SVM + Teknik Ensemble Bagging

Pengujian selanjutnya dengan menggunakan SVM dan salah satu teknik *ensemble* yaitu Bagging. Hasil pengujiannya sebagai berikut:

Tabel 7. Pengujian SVM + Teknik Ensemble Bagging

	Poly Kernel	Normalized Poly Kernel	RBF Kernel
Akurasi	91.04	82.08	56.71
TP Rate	0.91	0.821	0.567
FP Rate	0.087	0.204	0.567
Precision	0.91	0.832	0
Recall	0.91	0.821	0.567
F-Measure	0.908	0.8	0
MCC	0.845	0.695	0
ROC Area	0.954	0.903	0.593
PRC Area	0.927	0.834	0.491

Sumber: (Sugara & Subekti, 2019)

Dari Tabel 7. Pengujian SVM dengan salah satu teknik *ensemble* yaitu Bagging. Sama seperti teknik *ensemble* AdaBoost, *Poly Kernel* memiliki nilai akurasi terbaik sebesar 91.04% dengan TP Rate sebesar 0.91 dan FP Rate sebesar 0.087. Teknik *Esemble* Bagging ini juga menunjukkan dapat meningkatkan nilai akurasi terhadap perhitungan SVM menggunakan *poly kernel*.

KESIMPULAN

Pengujian pada penelitian deteksi dini gangguan autisme ini mengusulkan algoritma *support vector machine* (SVM) untuk memberikan nilai akurasi yang terbaik dengan menggunakan *small dataset*. Dataset yang dipakai pada pengujian ini sebanyak 67 dengan menghasilkan nilai akurasi yang tertinggi sebesar 85% pada *normalized poly kernel*. Dua teknik *ensemble* yaitu AdaBoost dan Bagging juga diusulkan dalam pengujian penelitian deteksi dini gangguan autisme ini untuk meningkatkan kinerja klasifikasi algoritma *support vector machine* (SVM). Berdasarkan hasil eksperimen yang telah dilakukan menunjukkan bahwa teknik *ensemble* menunjukkan performa dapat meningkatkan nilai akurasi. Model SVM dengan *poly kernel* dan teknik *ensemble* Bagging menunjukkan nilai akurasi tertinggi yaitu sebesar 91%.

REFERENSI

- Alarifi, H. S., & Young, G. S. (2018). Using Multiple Machine Learning Algorithms to Predict Autism in Children. *International Conference Artificial Intelligence*, 464–467.
- Bekerom, B. Van Den. (2012). Using Machine Learning for Detection of Autism Spectrum Disorder.
- Boham, S. E. (2013). Pola Komunikasi Orang Tua

- Dengan Anak Autis (Studi pada orang tua dari anak autis di Sekolah Luar Biasa AGCA Center Pumorow Kelurahan Banjer Manado). *Journal*, 11(4).
- Budiman, E., Santoso, E., & Afrianto, T. (2017). Pendeteksi Jenis Autis pada Anak Usia Dini Menggunakan Metode Linear Discriminant Analysis (LDA). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 1(August), 583-592.
- Feng, S., Zhou, H., & Dong, H. (2019). Using deep neural network with small dataset to predict material defects. *Materials and Design*, 162, 300-310.
<https://doi.org/10.1016/j.matdes.2018.11.060>
- Fridayanthie, E. W. (2015). Analisa Data Mining Untuk Prediksi Penyakit Hepatitis Dengan Menggunakan Metode Naive Bayes dan Support Vector Machine. *Jurnal Khatulistiwa Informatika*, 3, 24-36.
- Gardenia, M., Tursina, & Pratiwi, H. S. (2015). Sistem Pakar Deteksi Autisme Pada Anak Menggunakan Metode Fuzzy Tsukamoto. *Jurnal Sistem Dan Teknologi Informasi*, 2(1), 1-6.
- Hendita, G., Kusuma, A., Si, S., Cs, M., & Oktana, L. (2012). Sistem Identifikasi Penyakit Autis Anak Berbasis Web. *Jurnal TICOM*, 1(1), 29-41.
- Merianto, R. W. (2016). Peran Orang Tua Dalam Menangani Anak Autis. *Jom Fisip*, 3(1), 1-15.
- Munawarah, R., Soesanto, O., & Faisal, M. R. (2016). Penerapan Metode Support Vector Machine, 04(01), 103-113.
- Nurmasani, A., Utami, E., & Al Fatta, H. (2017). Analisis Support Vector Machine. *Jurnal Informasi Interaktif*, 2(1).
- Parapat, I. M., & Furqon, M. T. (2018). Penerapan Metode Support Vector Machine (SVM) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIHK) Universitas Brawijaya*, 2(10), 3163-3169.
- Permono, H. (2013). Peran Orang Tua dalam Optimalisasi Tumbuh Kembang Anak untuk Membangun Karakter Anak Usia Dini. *Prosiding Seminar Nasional Parenting*, 34-47.
- Puspitasari, A. M., Ratnawati, D. E., & Widodo, A. W. (2018). Klasifikasi Penyakit Gigi Dan Mulut Menggunakan Metode Support Vector Machine. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(2), 802-810.
- Rachman, F., & Purnami, W. S. (2012). Perbandingan Klasifikasi Tingkat Keganasan Breast Cancer Dengan Menggunakan Regresi Logistik Ordinal Dan Support Vector Machine (SVM). *Jurnal Sains Dan Seni ITS*, 1(1).
<https://doi.org/10.12962/j.23373520.v1i1.1932>
- Shaikhina, T., Lowe, D., Daga, S., Briggs, D., Higgins, R., & Khovanova, N. (2015). Machine learning for predictive modelling based on small data in biomedical engineering. *IFAC-PapersOnLine*, 28(20), 469-474.
<https://doi.org/10.1016/j.ifacol.2015.10.185>
- Sharma, S., & Sharma, V. (2016). Performance of Various Machine Learning Classifiers on Small Datasets with Varying Dimensionalities: A Study. *Circulation in Computer Science*, 1(1), 30-35.
<https://doi.org/10.22632/ccs-2016-251-23>
- Siwi, A. R. K., & Anganti, N. R. N. (2017). Strategi Pengajaran Interaksi Sosial pada Anak Autis. *Indigenous: Jurnal Ilmiah Psikologi*, 2(2), 184-192.
<https://doi.org/10.23917/indigenous.v2i2.5703>
- Sugara, Widyatmoko, Prakoso, S. D. (2018). Penerapan Algoritma C4.5 untuk Deteksi Dini Gangguan Autisme Pada Anak. *Seminar Nasional Teknologi Informasi Dan Komunikasi (SENTIKA)*, 87-96.
- Sugara, B., & Subekti, A. (2019). Penerapan Support Vector Machine (SVM) Pada Small Dataset Untuk Deteksi Dini Gangguan Autisme. *Laporan Akhir*.
- Susanto, S. E. (2014). Penerimaan Orang tua terhadap Kondisi Anaknya yang Menyandang Autisme di Rumah Terapis Little Star. *Jurnal Psikososial*, 9(2), 140-152.
- Susilowati, E., Sabariah, M. K., & Gozali, A. A. (2015). Implementasi Metode Support Vector Machine untuk Melakukan Klasifikasi Kemacetan Lalu Lintas Pada Twitter. *E-Proceeding of Engineering*, 2(1), 1-7.