

**“PENERAPAN METODE NAIVE BAYES DAN SUPPORT VECTOR
MACHINE (SVM) UNTUK KLASIFIKASI PENYAKIT
AUTISM SPECTRUM DISORDERS”**

Muhammad Eko Puji Listianto
(190441100078)

Pembimbing 1 : Dr. Cucun Very Angkoso, S.T., M.T.

Pembimbing 2 : Eza Rahmanita, S.T., M.T.

ABSTRAK

Autism Spectrum Disorder (ASD) adalah suatu gangguan perilaku seseorang yang disebabkan oleh kelainan struktural otak manusia yang mempengaruhi sistem saraf ditandai dengan adanya kendala dalam upaya untuk melakukan interaksi sosial, berkomunikasi serta aktifitas yang dilakukan secara berulang. Gangguan autisme biasanya didiagnosis di lingkungan klinis oleh spesialis yang sudah berlisensi dalam bidangnya dengan menggunakan prosedur yang bisa memakan waktu lama dan tidak hemat biaya. Oleh karena itu para ahli di bidang medis, psikologi dan ilmu perilaku terapan dalam beberapa dekade terakhir mengembangkan metode skrining seperti *Autism Spectrum Quetiont*. Namun metode skrining dinilai masih menggunakan perhitungan sederhana dengan menghitung skor jawaban akhir yang diberikan. Tentunya masih dibutuhkan suatu teknologi cerdas yang dapat menyelesaikan hal ini dengan cepat dan hemat biaya seperti *machine learning*. Pada penelitian ini menerapkan 2 metode yaitu *Naïve Bayes* dan *Support Vector Machine* untuk melakukan klasifikasi pada *Autism Spectrum Disorder*. Hasil penelitian ini menunjukkan bahwa baik metode *Naïve Bayes* maupun *Support Vector Machine* efektif digunakan dalam klasifikasi penyakit *Autism Spectrum Disorders* (ASD). Metode *Naïve Bayes* mencapai akurasi sebesar 97% dengan nilai *recall (sensitifity)* sebesar 97% serta nilai *specificity* mencapai 97% sedangkan Metode SVM bahkan mampu mencapai akurasi sebesar 100% dengan nilai *recall (sensitifity)* dan *specificity* juga mencapai 100%.

Kata Kunci : Klasifikasi, *Autism Spectrum Disorder*, *Naïve Bayes*, *Support Vector Machine*

**“IMPLEMENTATION OF NAIVE BAYES AND SUPPORT VECTOR
MACHINE (SVM) METHODS FOR AUTISM SPECTRUM
DISORDERS CLASSIFICATION”**

Muhammad Eko Puji Listianto
(190441100078)

Pembimbing 1 : Dr. Cucun Very Angkoso, S.T., M.T.

Pembimbing 2 : Eza Rahmanita, S.T., M.T.

ABSTRACT

Autism Spectrum Disorder (ASD) is a behavioral disorder caused by structural abnormalities in the human brain that affect the nervous system, characterized by challenges in social interaction, communication, and repetitive activities. The diagnosis of autism is typically conducted in clinical settings by licensed specialists using time-consuming and costly procedures. In response to this, experts in the fields of medicine, psychology, and applied behavioral science have developed screening methods such as the Autism Spectrum Questionnaire over the last few decades. However, these screening methods are considered to rely on simplistic calculations by computing the final answer score. Therefore, there is a need for intelligent technology, specifically machine learning, to address this efficiently and cost-effectively. This study applies two methods, Naïve Bayes and Support Vector Machine, for the classification of Autism Spectrum Disorder. The results indicate that both Naïve Bayes and Support Vector Machine methods are effective in classifying Autism Spectrum Disorders (ASD). Naïve Bayes achieved an accuracy of 97% with recall (sensitivity) and specificity both reaching 97%. Meanwhile, the SVM method achieved an impressive accuracy of 100%, with recall and specificity also reaching 100%.

Keywords: *Classification, Autism Spectrum Disorder, Naïve Bayes, Support Vector Machine*

DAFTAR ISI

HALAMAN JUDUL.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN ORISINALITAS.....	iv
MOTTO	v
ABSTRAK	vi
KATA PENGANTAR	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL.....	xiii
KODE PROGRAM.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.2.1 Permasalahan.....	4
1.2.2 Metode Usulan	4
1.2.3 Pertanyaan Penelitian	4
1.3 Tujuan dan Manfaat.....	4
1.3.1 Tujuan Penelitian.....	4
1.3.2 Manfaat Penelitian.....	5
1.4 Batasan Masalah	5
1.5 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA.....	7
2.1 Landasan Teori	7
2.1.1 <i>Autism Spectrum Disorder</i>	7

2.1.2 <i>Data mining</i>	8
2.1.3 Machine Learning.....	10
2.1.4 Klasifikasi.....	10
2.1.5 <i>Standard Scaler</i>	11
2.1.6 <i>Naïve Bayes</i>	12
2.1.7 <i>Support Vector Machine (SVM)</i>	14
2.1.8 Kernel pada <i>Support Vector Machine (SVM)</i>	16
2.1.9 <i>K Fold Cross Validation</i>	21
2.1.10 <i>Confusion Matrix</i>	22
2.2 Penelitian Terkait.....	24
BAB III METODE PENELITIAN.....	33
3.1 Business Understanding.....	33
3.2 Data Understanding	34
3.3 Data Preparation	35
3.3.1 <i>Data Cleaning</i>	35
3.3.2 <i>Data Selection</i>	36
3.3.3 <i>Data Transformation</i>	38
3.4 Modelling.....	40
3.5 Evaluation	45
3.6 Deployment.....	45
BAB IV HASIL DAN PEMBAHASAN	46
4.1 Data Understanding	46
4.2 Data Preparation	50
4.2.1 <i>Data Cleaning</i>	50
4.2.2 <i>Data Selection</i>	52
4.2.3 Data Transformation.....	53

4.3 Modeling.....	55
4.4 Evaluation	59
4.4.1 Skenario Pengujian.....	59
4.5 Deployment.....	63
4.5.1 Instalasi streamlit.....	64
4.5.2 Kode aplikasi Streamlit	64
4.5.3 Menjalankan aplikasi Streamlit	71
4.5.4 Tampilan aplikasi web.....	72
BAB V PENUTUP.....	75
5.1 Kesimpulan	75
5.2 Saran	75
DAFTAR PUSTAKA	76
LAMPIRAN.....	80
DAFTAR RIWAYAT HIDUP	81

DAFTAR GAMBAR

Gambar 2. 1 Tipe Kernel dalam SVM	14
Gambar 2. 2 Konsep cara kerja pada SVM.....	15
Gambar 2. 3 Gambaran ketika parameter C dengan nilai besar.....	18
Gambar 2. 4 Gambaran ketika parameter C dengan nilai rendah	18
Gambar 2. 5 Gambaran ketika parameter gamma dengan nilai tinggi.....	18
Gambar 2. 6 Gambaran ketika parameter gamma dengan nilai rendah	19
Gambar 2. 7 K-fold Cross Validation	22
Gambar 3. 1 Cross Industry Standart Process for Data Mining.....	33
Gambar 3. 2 Sampel data	36
Gambar 3. 3 Tipe data sebelum dilakukan label encoding	39
Gambar 3. 4 Tipe data setelah dilakukan label encoding.....	40
Gambar 3. 5 Diagram alur proses klasifikasi model	41
Gambar 3. 6 Flowchart tahapan metode Gaussian Naive Bayes	42
Gambar 3. 7 Flowchart tahapan metode SVM.....	43
Gambar 3. 8 Skenario Pengujian.....	44
Gambar 4. 1 Jenis tipe data atribut pada dataset	49
Gambar 4. 2 Jumlah missing value pada tiap atribut	51
Gambar 4. 3 Penanganan missing value dengan mean	52
Gambar 4. 4 Jenis tipe data atribut setelah penanganan missing value	53
Gambar 4. 5 Jenis tipe data tiap atribut.....	54
Gambar 4. 6 Confusion Matrix Gaussian Naive Bayes	59
Gambar 4. 7 Confusion Matrik Support Vector Machine.....	61
Gambar 4. 8 Tampilan web app autisme.....	73

DAFTAR TABEL

Tabel 2. 1 Confusion Matrix	23
Tabel 2. 2 Penelitian terkait	28
Tabel 3. 1 Keterangan atribut atau variabel data.....	36
Tabel 3. 2 Penjelasan pertanyaan proses skrining.....	37
Tabel 4. 1 Nilai koefisien korelasi tiap atribut.....	47
Tabel 4. 2 Sampel dataset penelitian.....	47
Tabel 4. 3 Keterangan atribut pada dataset	48
Tabel 4. 4 Penggunaan kernel pada SVM.....	55
Tabel 4. 5 Parameter dalam K-Fold Cross Validation	56
Tabel 4. 6 Classification Report Gaussian Naive Bayes.....	60
Tabel 4. 7 Classification Report Support Vector Machine	62
Tabel 4. 8 Hasil classification report seluruh skenario	63

KODE PROGRAM

Kode Program 4. 1 kode informasi data.....	49
Kode Program 4. 2 Kode mencari nilai missing value.....	50
Kode Program 4. 3 Kode mengatasi missing value dengan mean	51
Kode Program 4. 4 Mengubah kategorikal menjadi numerik Label Encoder.....	53
Kode Program 4. 5 Informasi tipe data setiap kolom.....	54
Kode Program 4. 6 Proses modeling sampai dengan evaluasi kinerja.....	57
Kode Program 4. 7 Install streamlit	64
Kode Program 4. 8 Install numpy pandas	64
Kode Program 4. 9 Menyimpan model dalam format pickle.....	64
Kode Program 4. 10 Impor model format pickle	65
Kode Program 4. 11 Kode implementasi aplikasi streamlit.....	70
Kode Program 4. 13 Menjalankan aplikasi web streamlit	71