

CLASSIFICATION OF UNDERSTANDING OF STUDENTS IN LEARNING KITAB
KUNING USING SUPPORT VECTOR MACHINE
(CASE STUDY: ISLAMIC BOARDING SCHOOL MAMBAUS SHOLIHIN)
KLASIFIKASI PEMAHAMAN SANTRI DALAM PEMBELAJARAN KITAB KUNING
MENGUNAKAN METODE SUPPORT VECTOR MACHINE
(STUDI KASUS: PONDOK PESANTREN MAMBAUS SHOLIHIN)
Nabila Nur Amalia¹, Bernadus Anggo Seno Aji², Regita Putri Permata³

¹²³ Teknologi Informasi, Institut Teknologi Telkom Surabaya, Indonesia

¹nabill_nna@student.ittelkom-sby.ac.id, ²Bernadus.seno@ittelkom-sby.ac.id,
³regitapermata@ittelkom-sby.ac.id

Informasi Artikel

Received: December 2020

Revised: January 2021

Accepted: January 2021[1]

Published: February 2021

Menggunakan style **info**

Abstract (menggunakan style abstract)

Objective: To apply and test the performance of the classification system at the Mambaus Sholihin Islamic Boarding School so that it is more effective in assessing students' understanding of benchmarks based on the grades obtained while studying.

Design/method/approach: Using the Support Vector Machine algorithm in system training and testing with kernel functions and parameters, To compare the results of accuracy, precision, and recall.

Results: Based on the kernel function comparison performance test, the Gaussian-RBF (Radial Basis Function) kernel is the most suitable kernel function to produce the best predictions. The best accuracy results were obtained by the RBF kernel, which obtained an accuracy of 73% for training data and 89% for testing data.

Authenticity/state of the art: This study has differences in the methods used for classifying understanding of the Kitab Kuning using the SVM method and the application of trick kernels, which has never been done in previous studies.

Abstrak (menggunakan style abstrak)

Keywords: Arabic, Classification, Kitab Kuning, Nahwu, Report Card Value, SVM Algorithm.

Kata kunci: Algoritma SVM, Bahasa Arab, Kitab Kuning, Klasifikasi, Nahwu, Nilai Raport.

Tujuan: Untuk menerapkan dan menguji performansi sistem klasifikasi di Pondok Pesantren Mambaus Sholihin agar lebih efektif dalam menilai tolak ukur pemahaman santri berdasarkan nilai yang didapatkan saat menempuh pendidikan. Untuk mengimplementasikan sistem klasifikasi santri dalam memahami kitab kuning menggunakan metode *Support Vector Machine* dengan website.

Perancangan/metode/pendekatan: Menggunakan algoritma *Support Vector Machine* dalam pelatihan dan pengujian sistem dengan fungsi kernel dan parameter, Untuk membandingkan hasil akurasi, presisi, dan recall.

Hasil: Berdasarkan uji performa perbandingan fungsi kernel, kernel *Gaussian-RBF (Radial Basis Function)* merupakan fungsi kernel yang paling cocok untuk menghasilkan prediksi paling baik. Hasil akurasi terbaik didapatkan oleh kernel *RBF*, yakni mendapatkan akurasi sebesar 73% untuk data training dan untuk data testing sebesar 89%.

Keaslian/ *state of the art*: Penelitian ini memiliki perbedaan dalam metode yang digunakan untuk klasifikasi pemahaman terhadap kitab kuning dengan menggunakan metode SVM dan penerapan kernel trik, yang mana belum pernah dilakukan pada penelitian-penelitian terdahulu.

1. Pendahuluan

Pentingnya Pendidikan dalam kehidupan tidak dapat diabaikan. Pondok Pesantren memiliki peran yang sangat penting dalam memajukan, mengembangkan, memberi pengetahuan, serta membentuk karakter pada diri masing-masing. Peran serta keberadaan Pondok Pesantren pun memiliki nilai yang sangat penting dalam kemajuan sektor Pendidikan, terutama di Indonesia[1].

Pondok Pesantren Mambaus Sholihin ini berada di kota Gresik, Jawa Timur. Kondisi saat ini pada Pondok Pesantren Mambaus Sholihin untuk Menentukan santri itu dikatakan lulus untuk wisuda dan dinyatakan paham atau tidak dalam pembelajaran kitab kuning maka diadakan seleksi 2 tahap. Tahap pertama yakni Ujian UTS dan UAS untuk mendapatkan nilai dan Tahap Kedua yakni dengan melakukan tes lisan hafalan kitab *Alfiyah Ibnu Malik* dan membaca kitab kuning gundul. Namun, terdapat kekurangan dari dua tahap tersebut, yakni membutuhkan lebih banyak waktu untuk melakukan proses seleksi 2 tahap dengan jumlah santri berkisar antara 300-500 orang dengan jumlah penguji berkisar 20-50 orang, juga metode *SVM* ini mencegah data yang terbilang subjektif.

Untuk mengatasi kekurangan tersebut, maka perlunya sistem baru yang berisikan Klasifikasi untuk menentukan apakah santri tersebut dapat dinyatakan paham dan tidak paham dalam mempelajari kitab kuning. Berdasarkan penelitian sebelumnya yang dilakukan oleh Cep Lukman Rohmat, Irfan Ali, Aria Pratama, Farid Ali Ma'ruf, 2021, "Analisa Pemahaman santri dalam mempelajari Kitab Kuning Menggunakan Metode *Algoritma Naïve Bayes* di Pondok Al Arifah Buntet Pesantren Cirebon" dengan hasil akurasi sebesar 89%[2].

Selain itu, penelitian Ahmad Hudawi AS, Nur Octavia, Andik Elfandiono, Andi Bayu Setiawan, Alfian Abdul Ghafur, Aldiniyah Eko Susanto, 2021, "Klasifikasi Pemahaman Santri dalam Pembelajaran Kitab Kuning menggunakan Algoritma C4.5", Algoritma C4.5 berbasis Forward Selection berhasil mendapatkan akurasi lebih tinggi, mencapai angka 85,68%[3].

Berdasarkan penelitian sebelumnya, diputuskan untuk menggunakan metode Algoritma *Support Vector Machine (SVM)* pada penelitian ini. Algoritma *Support Vector Machine (SVM)* dipilih karena memiliki karakteristik yang kuat dan cenderung lebih akurat, serta dianggap sesuai dan tepat dengan atribut-atribut yang telah ditetapkan. Selain itu, SVM mempunyai salah satu kelebihan yaitu penanganan kesalahan pada set data latih menggunakan *Structural Risk Minimization (SRM)*. SRM dianggap sebagai metode yang baik karena tidak hanya mengurangi kesalahan yang terjadi, namun juga mempertimbangkan faktor lain yang berperan[4].

Algoritma ini telah terbukti memiliki kekuatan dan ketepatan dalam menghasilkan klasifikasi yang tepat. Dibuktikan dengan beberapa penelitian sebelumnya, seperti pada penelitian Fatin Hilmiyah, 2017, “Prediksi Kinerja Mahasiswa menggunakan *Support Vector Machine* menggunakan pengelola Program Studi Di Perguruan Tinggi (Studi Kasus: Program Studi Magister Statistika ITS)”, dari hasil uji coba Algoritma SVM mampu mencapai tingkat akurasi yang lebih tinggi mencapai 74,84%, presisi sebesar 81,23%, dan *recall* sebesar 90,49% [5].

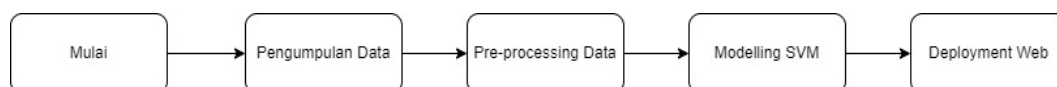
Demikian juga, penelitian Jefry Pranata M, 2018, “Klasifikasi kabupaten di provinsi Jawa Timur berdasarkan indikator daerah tertinggal dengan metode *Support Vector Machine* dan *Entropy Based Fuzzy Support Vector Machine*” dengan hasil akurasi sebesar 86,61% dengan penerapan kernel *RBF* dan *Linear*[6].

Tidak hanya itu, penelitian Dea Andia Rachmawati, 2016, “Klasifikasi Teks Permintaan Informasi untuk aplikasi online shop menggunakan Algoritma *Support Vector Machine* (Studi Kasus: Bento Shop)” dengan hasil akurasi sebesar 94,74%, presisi 93,18%, *recall* 96,09% dengan penerapan kernel *linear* dan *RBF*[7].

Berdasarkan penelitian terdahulu, metode yang akan digunakan adalah Algoritma *Support Vector Machine* dengan empat kernel trik yaitu *Linear*, *Radial Basis Function (RBF)*, *Polynomial* dan *Sigmoid*. Dengan menggunakan kernel trik maka dapat diketahui mana kernel yang menghasilkan akurasi paling optimal pada klasifikasi.

2. Metode Penelitian

Metode penelitian yang digunakan pada penelitian ini meliputi seperti pada Gambar 1. Yang mana metode penelitian diawali dengan Pengumpulan Data, *Preprocessing* data, modelling dengan metode *Support Vector Machine* dan *deployment* web.



Gambar 1. Metode Penelitian

- a) Tahap pertama yaitu proses pengumpulan data. Pengumpulan data dilakukan dengan cara mendatangi Madrasah Aliyah Mambaus Sholihin dengan membawa surat pengantar pengambilan data. Kemudian data didapatkan dengan beberapa kali kunjungan.
- b) Data yang telah didapatkan sesuai kebutuhan disimpan dan diproses dalam tahap *Pre-processing* data.
- c) Tahapan selanjutnya adalah *Modelling SVM*. Pada tahapan ini data diolah dengan menggunakan metode *Support Vector Machine* melalui beberapa tahap yang nantinya

akan dijelaskan pada tahap metode *Support Vector Machine*.

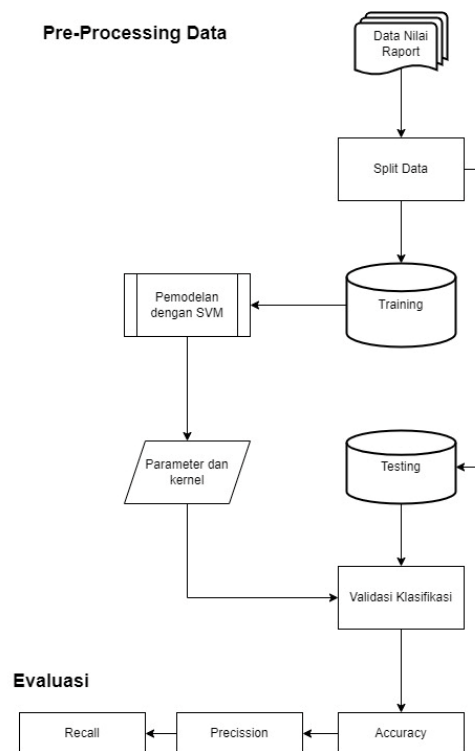
- d) Tahapan yang terakhir adalah *deployment* web. Pada tahapan ini memilih framework yang akan digunakan dan digabungkan dengan model *Support Vector Machine* yang telah diperoleh. Hasil klasifikasi akan ditampilkan dalam web tersebut.

2.1 Data Penelitian

Pada penelitian ini menggunakan metode pengumpulan data *sekunder*. Untuk data yang digunakan pada penelitian ini adalah data akademik santri pada Madrasah Aliyah Mambaus Sholihin yang berjumlah 280 data santri. Dengan data training yang digunakan sejumlah 70% data sejumlah 190 data dan Data Testing yang digunakan sejumlah 30% data selain Data Training yaitu sebanyak 80 data. Data tersebut meliputi beberapa atribut yakni: Nilai Raport Nahwu, juga Nilai UTS dan UAS, Nilai Raport Bahasa Arab, Nilai Baca Kitab, Jenjang Waktu menjadi Santri.

2.2 Pre-processing Data

Sebelum melakukan modelling dengan metode *SVM*, maka dilakukan *preprocessing* data terlebih dahulu[5]. Berikut adalah Diagram alir *Preprocessing* data:



Gambar 2. Pre-processing Data

Pengolahan data dimulai dengan langkah *Preprocessing* data. Pada tahap ini, dilakukan analisis distribusi data untuk memahami pola distribusi dalam variabel input dan output. Tahap berikutnya adalah pemodelan klasifikasi menggunakan algoritma *Support Vector Machine* (*SVM*). Kemudian data dibagi menjadi dua bagian melalui proses *Split data*, yakni data

pelatihan (*training*) dan data pengujian (*testing*). Tahap Selanjutnya adalah proses silang, dimana data pengujian digunakan sebagai data pelatihan sebelumnya dijadikan data pengujian. Setelah Data training siap diolah, melakukan pembuatan model klasifikasi dengan SVM untuk mendapatkan Parameter dan Kernel yang sesuai. Setelah mendapatkan model SVM untuk klasifikasi, model tersebut akan diimplementasikan pada data testing. Tahap terakhir adalah evaluasi dengan menggunakan matrik evaluasi seperti *Recall*, *Precision* dan *Akurasi* [5].

2.3 Algoritma Support Vector Machine

Algoritma *Support Vector Machine* (SVM) adalah bentuk algoritma *supervised* yang digunakan untuk klasifikasi dengan mempartisi data menjadi dua kelas melalui penggunaan garis vektor yang disebut *hyperplane*. Pada penelitian ini menggunakan teknik kernel trik. Berikut adalah persamaan dari parameter dan kernel yang digunakan:

a) Kernel *linear*

Fungsi kernel ini cocok diterapkan saat data yang sedang dianalisis dapat dipisahkan secara *linear*. Berikut adalah Persamaan fungsi Kernel *Linear*[6]:

$$K(x, xi) = (x * xi) \quad (1)$$

b) Kernel *RBF*

Fungsi kernel ini dikenal karena kemampuannya yang baik dalam performa, terutama ketika parameter-parameter tertentu dipilih dengan tepat. Berikut adalah persamaan fungsi Kernel *RBF*[6]:

$$K(x, xi) = \exp(-\gamma((x - xi)^2)), \gamma > 0 \quad (2)$$

c) Kernel *Polinomial*

Fungsi kernel *Polynomial* adalah fungsi kernel yang sangat sesuai digunakan dalam metode SVM dan kernelisasi lainnya, yang mana fungsi kernel *polynomial* juga terbukti efektif dalam menangani tugas klasifikasi pada dataset pelatihan yang telah dinormalisasi. Berikut adalah persamaan fungsi Kernel *Polynomial*[6]:

$$K(x, xi) = (\gamma x, xi + r)^p, \gamma > 0 \quad (3)$$

d) Kernel *Sigmoid*

Ketika jumlah fitur semakin banyak, nilai *gamma* yang optimal cenderung lebih kecil, sedangkan sebaliknya, jika jumlah fitur sedikit, nilai *gamma* yang optimal cenderung lebih besar. Berikut adalah persamaan fungsi Kernel *Sigmoid*[8]:

$$K(X, Y) = \tanh(\gamma X * Y + c) \quad (4)$$

e) Gamma

Hyperparameter gamma mengontrol pengaruh yang dimiliki pada setiap kasus pada posisi garis *Hyperplane* dan digunakan oleh semua fungsi kernel kecuali kernel *linear*.

Berikut persamaan SVM dengan *Gamma*[9]:

$$\sum(\alpha_i * y_i * K(x_i, x)) + b \quad (5)$$

Dengan persamaan tersebut dan meninjau beberapa referensi yang digunakan, Maka dalam penelitian ini, besarnya nilai *Gamma* yang digunakan adalah 0,001; 0,001; 0,1; 1, 10, dan 100[9].

f) Cost

Hyperparameter Cost atau *C* dalam *soft-margin SVM* menetapkan nilai *C* untuk memiliki *case* didalam *margin*. Berikut persamaan *SVM* dengan *Cost*[8]:

$$\sum(\alpha_i * y_i * K(x_i * x)) + b \quad (6)$$

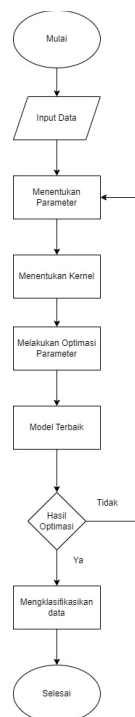
Dengan persamaan tersebut dan meninjau beberapa referensi yang digunakan, Maka dalam penelitian ini, nilai *Cost* yang digunakan adalah : 0,001; 0,1; 1, 10 dan 100.

g) Degree

Hyperparameter degree adalah parameter yang terkait dengan kernel *polynomial*. Parameter *degree* mengontrol tingkat dari *polinomial* yang digunakan. Berikut persamaan *SVM* dengan *Degree*[9]:

$$\sum(\alpha_i * y_i * (x_i \cdot x + 1)^d) + b \quad (7)$$

Dengan persamaan tersebut dan meninjau beberapa referensi yang digunakan, maka pada penelitian ini, nilai *degree* yang digunakan adalah 1, 2, dan 3.



Gambar 3. Flowchart Algoritma Support Vector

Berikut penjelasan Flowchart Algoritma *Support Vector Machine* pada **gambar 3** adalah:

a) **Input Data**

Memasukkan data yang telah dilakukan *preprocessing*. Adapun data yang digunakan berjumlah 270 yang telah dilakukan *preprocessing* dari total data 308.

b) Menentukan Parameter

Dalam penelitian ini menggunakan beberapa parameter dan kernel[9]. Yaitu sebagai berikut:

- *Linear*
Cost: 0,001; 0,1; 1; 10; 100
- *RBF*
Gamma: 0,0001; 0,001; 0,1; 1; 10; 100
- *Polynomial*
Cost: 0,1; 1; 10; 100
Degree: 1, 2, 3
- *Sigmoid*
Cost: 0,1; 1; 10; 100
Gamma: 0,0001; 0,001; 0,1; 1; 10; 100

c) Melakukan Optimasi Parameter

Mengoptimalkan setiap parameter yang telah dipilih, agar mendapatkan hasil yang sesuai dengan harapan. Setelah hasil optimasi dinilai baik, maka lanjut ke tahap selanjutnya. Saat hasil optimasi kurang baik maka kembali pada tahap awal.

d) Mendapatkan Model Terbaik

Setelah melakukan optimasi pada parameter, di dapatkanlah model terbaik. Model terbaik didapatkan dari percobaan sistem terhadap setiap parameter. Diambil dari akurasi, presisi dan recall terbaik.

e) Mengklasifikasikan Data

Mengklasifikasikan hasil yang telah didapatkan sesuai dengan target yang ingin dicapai. Klasifikasi dengan metode SVM yang dimana mengoptimalkannya dengan model kernel trik yang telah terpilih. Targetnya adalah paham, dan tidak paham.

2.4 Evaluasi Klasifikasi

Salah satu evaluasi klasifikasi yang mengindikasikan performa secara keseluruhan dari metode teknik klasifikasi disebut tingkat akurasi klasifikasi (Nugroho, dkk,2003). Semakin besar akurasi klasifikasi berarti menunjukkan performa dari metode teknik klasifikasi. Formula berikut dapat digunakan untuk menghitung evaluasi kinerja metode klasifikasi[10]:

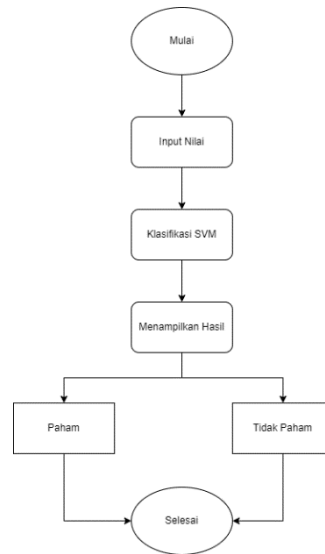
$$\text{Akurasi} = \frac{TN}{TN+TP+FN+FP} \times 100\% \quad (8)$$

$$\text{Presisi} = \frac{TP}{TP+FP} \times 100\% \quad (9)$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\% \quad (10)$$

2.5 Implementasi Web

Adapun Diagram alir untuk desain dan implementasi sistem pada web, dapat dilihat pada **Gambar 2**:



Gambar 4. Flowchart Aplikasi web dengan Algoritma SVM

Berikut untuk penjelasan diagram alir **Gambar 3** :

- Masukkan nilai dengan beberapa inputan, yaitu Nilai Raport Nahwu, Raport Bahasa Arab, Baca Kitab, Nahwu UAS, Nahwu UTS dan Jenjang waktu menjadi santri.
- Setelah nilai dimasukkan akan ada proses pengklasifikasian dengan target paham dan tidak paham.
- Kemudian hasil akan ditampilkan.
- Hasil yang ditampilkan berupa 2 pilihan yakni Paham dan Tidak paham
- Selesai

3 Hasil dan Pembahasan

3.1 Hasil Pengumpulan Data

Tabel 1. Sampel Data

Nahwu UTS	Nahwu UAS	Nahwu nilai raport	B.Arab Nilai Raport	Baca kitab	Jenjang waktu	Keterangan
35	56	62	23	22	3	Tidak paham
60	88	82	43	15	3	Paham
40	36	98	68	27	3	Tidak paham
25	80	96	45	24	3	Tidak paham
25	36	85	20	25	3	paham

Pada **tabel 1** adalah beberapa data dari data asli yang telah dilakukan tahap preprocessing. Adapun jumlah data pada target **keterangan** sebanyak 195 data santri yang dinyatakan Paham dan 85 data santri yang dinyatakan Tidak Paham.

3.2 Pelatihan dan Pengajuan Model

3.2.1 Uji Performa Fungsi Kernel Terbaik

Pelatihan dan pengujian model memiliki tujuan untuk mendapatkan fungsi kernel yang dapat dikatakan optimal agar bisa diterapkan dalam model klasifikasi dalam penelitian ini. Pada penelitian ini menggunakan perbandingan 4 kernel SVM yaitu Kernel *Linear*, Kernel *Gaussian-RBF (Radial Basis Function)*, Kernel *Polinomial* dan Kernel *Sigmoid*. Optimasi parameter dilakukan dengan menggunakan metode *Grid Search*. Dengan rentang *gamma* yang digunakan sebagai berikut : 0,0001; 0,001; 0,1; 1; 10; 100; juga dengan rentang nilai Cost : 0,001; 0,1; 1; 10; 100; juga nilai *degree*: 1, 2, 3. **Tabel 2** adalah hasil akurasi dari uji coba implementasi SVM dengan kernel yang digunakan.

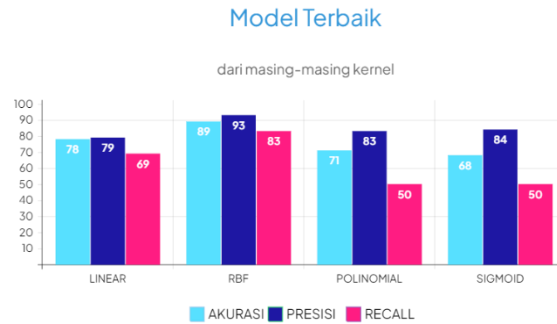
Tabel 2. Hasil akurasi dari Pemodelan SVM

No	Kernel	Hiperparameter			Akurasi	Presisi	Recall
		C	Gamma	Degree			
1							
2	Linear	0.001			0,74	0,72	0,61
3	Linear	0.1			0,77	0,77	0,67
4	Linear	1			0,78	0,79	0,69
5	Linear	10			0,78	0,79	0,69
6	Linear	100			0,78	0,79	0,69
7	RBF		0,0001		0,68	0,34	0,50
8	RBF		0,001		0,89	0,93	0,83
9	RBF		0.1		0,71	0,85	0,50
10	RBF		1		0,71	0,85	0,50
11	RBF		10		0,71	0,85	0,50
12	RBF		100		0,71	0,85	0,50
13	Polinomial	1		1	0,71	0,85	0,50
14	Polinomial	10		2	0,70	0,75	0,49
15	Polinomial	100		3	0,66	0,48	0,50
16	Sigmoid	0.1	0.1		0,68	0,84	0,50
17	Sigmoid	1	1		0,68	0,84	0,50
18	Sigmoid	10	0,0001		0,52	0,38	0,41
19	Sigmoid	100	0,0001		0,46	0,37	0,37
20	Sigmoid	100	100		0,68	0,84	0,50

Pada **tabel 2** terlihat masing-masing akurasi pada setiap hiperparameter yang telah ditentukan. Didapatkan model terbaik menggunakan kernel RBF dengan akurasi sebesar 89%, presisi sebesar 93% dan recall sebesar 83%.

3.2.2 Pengaruh Kernel yang digunakan dalam SVM

Kernel yang digunakan pada penelitian ini adalah kernel *Linear*, kernel *RBF*, kernel *Polinomial* dan kernel *Sigmoid*. Dari keempat kernel, akan didapatkan satu kernel terbaik yang memisahkan kedua buah class. Berikut merupakan grafik yang dihasilkan menggunakan kernel *Linear*, *RBF*, *Polinomial* dan *Sigmoid*.

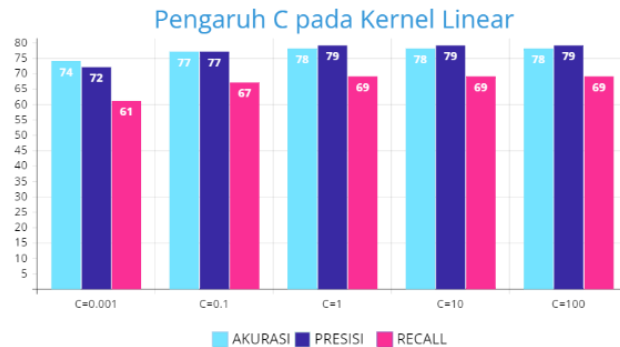


Gambar 5. Diagram batang model terbaik

Dan pada **gambar 5** adalah grafik hasil dari model terbaik masing-masing kernel. Terlihat akurasi tertinggi ada pada kernel *RBF* dengan jumlah akurasi sebanyak 89%. Dan akurasi terendah ada pada kernel *Sigmoid* dengan akurasi sebesar 68%. Kemudian untuk jumlah akurasi pada kernel *Linear* sebesar 78%, juga pada kernel *Polinomial* sebesar 71% dan pada kernel *Sigmoid* sebesar 68%.

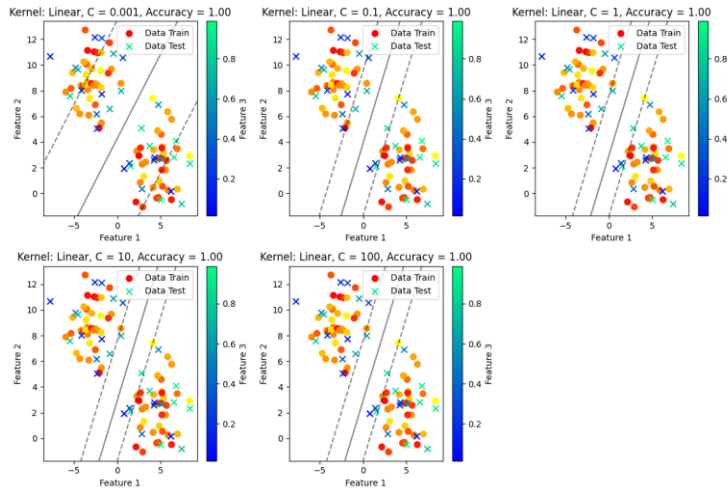
3.2.3 Pengaruh parameter C (Cost) pada kernel Linear

Berikut perbandingan nilai parameter d dengan nilai C pada kernel *linear* berdasarkan akurasi yang dihasilkan:



Gambar 6. Pengaruh C pada Linear

Grafik pada **gambar 6** menunjukkan bahwa semakin besar nilai parameter C yang digunakan pada kernel *Linear* belum tentu menghasilkan akurasi terbaik dalam sebuah pemodelan. Nilai parameter C pada penelitian ini sudah optimal pada $C=1$. Akurasi yang dihasilkan pada nilai $C=1$ yaitu 78%, juga pada nilai $C=10$ dan $C=100$. Nilai C default yang digunakan sendiri $C=1$.



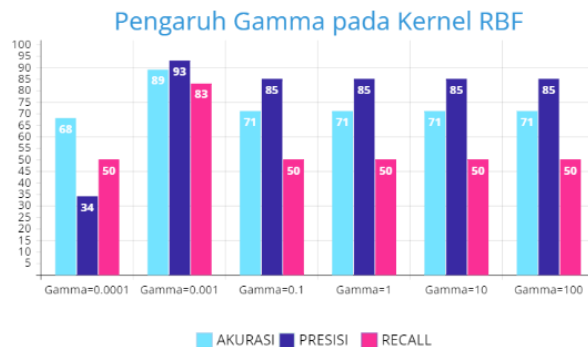
Gambar 7. Ilustrasi pada Kernel *Linear*

Pada **gambar 7** adalah ilustrasi pada kernel *linear* dengan parameter *C*. Dalam percobaan ini model terbaiknya yaitu $C=1$ dengan *feature 1* (atribut nilai baca kitab), *feature 2* (atribut nilai nahwu raport), *feature 3* (atribut nilai b.arab raport) dengan data uji (plot bulat berwarna merah) dan data latih (plot x berwarna hijau muda). Hasil ilustrasi adalah gambar garis solid miring dengan garis nonlinear yang berjarak dekat, dengan sebagian besar plot bulat berwarna merah dan kuning berada disisi kosong garis pemisah nonlinear.

Berdasarkan hasil ilustrasi tersebut diambil kesimpulan bahwa semakin kecil nilai *C* maka besar kemungkinan terjadinya kesalahan klasifikasi, dan semakin besar nilai *C* maka semakin kecil kemungkinan terjadinya kesalahan klasifikasi.

3.2.4 Pengaruh parameter Gamma pada kernel RBF

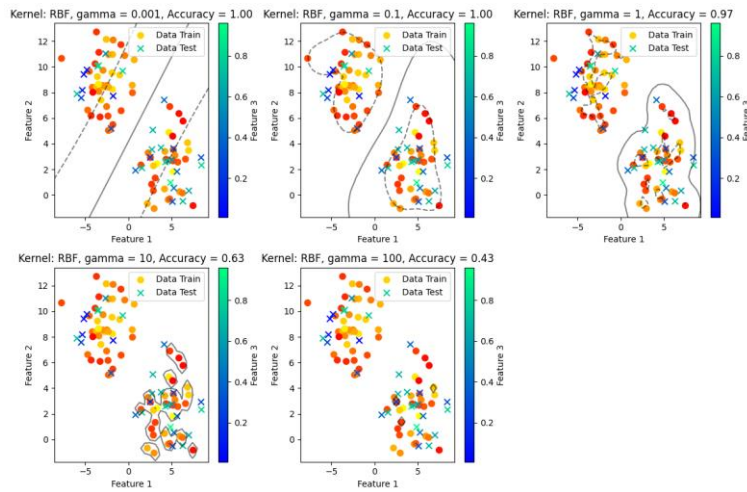
Berikut perbandingan nilai parameter *Gamma* pada kernel *RBF* berdasarkan akurasi yang dihasilkan:



Gambar 8. Pengaruh Gamma pada kernel RBF

Grafik pada **gambar 8** menunjukkan bahwa semakin besar nilai parameter *Gamma* yang digunakan pada kernel *RBF* belum tentu menghasilkan akurasi terbaik dalam sebuah pemodelan, dan juga menghasilkan nilai yang berbeda. Nilai parameter *Gamma* pada penelitian ini sudah optimal pada $Gamma=0,001$. Akurasi yang dihasilkan pada nilai $Gamma=0,001$ yaitu

89% dan menjadi pemodelan terbaik pada penelitian ini. Nilai *Gamma* default yang digunakan sendiri $\Gamma=1$.



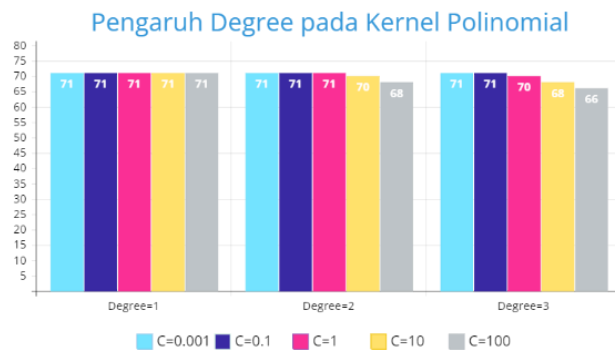
Gambar 9. Ilustrasi pada Kernel RBF

Pada **gambar 9** adalah ilustrasi pada kernel *RBF* dengan parameter *Gamma*. Dalam percobaan ini model terbaiknya yaitu $\Gamma=0,001$ dengan *feature 1* (atribut nilai baca kitab), *feature 2* (atribut nilai nahwu raport), *feature 3* (atribut nilai b.arab raport) dengan data uji (plot bulat berwarna merah) dan data latih (plot x berwarna hijau muda). Hasil ilustrasi adalah gambar garis solid miring dengan garis pemisah nonlinear yang berjarak..

Berdasarkan hasil ilustrasi, dapat diambil kesimpulan bahwa semakin kecil nilai *Gamma* maka semakin kecil risiko kemungkinan terjadinya kesalahan klasifikasi, dan semakin besar nilai *Gamma* maka semakin banyak kemungkinan terjadi kesalahan klasifikasi.

3.2.5 Pengaruh parameter Degree pada kernel Polinomial

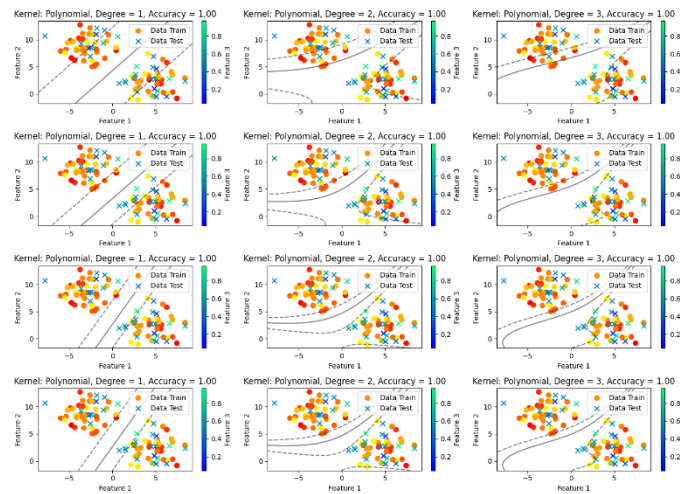
Berikut perbandingan nilai parameter *d* dengan nilai *C* pada kernel *Polinomial* berdasarkan akurasi yang dihasilkan:



Gambar 10. Pengaruh Degree pada kernel Polinomial

Grafik pada **gambar 10** menunjukkan bahwa semakin besar nilai parameter *d* yang digunakan pada kernel *Polinomial* belum tentu menghasilkan akurasi terbaik dalam sebuah pemodelan, dan juga menghasilkan nilai akurasi yang cenderung sama. Nilai parameter *d* pada penelitian ini

paling optimal pada $d=1$. Akurasi yang dihasilkan pada nilai $d=1$ dan $C=0,001$ yaitu 71% pada penelitian ini. Nilai *degree* default yang digunakan sendiri $d=3$.



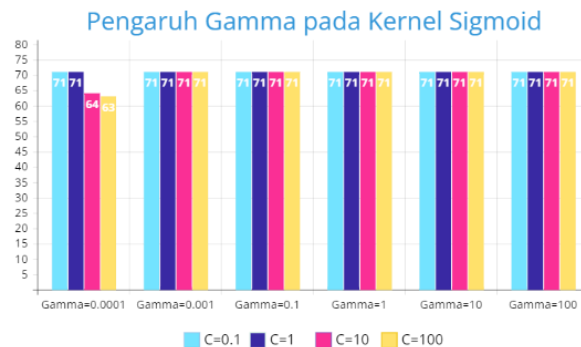
Gambar 11. Ilustrasi pada Kernel Polinomial

Pada **gambar 11** adalah ilustrasi pada kernel *Polinomial* dengan parameter C dan *degree*. Dalam percobaan ini model terbaiknya yaitu $C=1$ dan *degree*=1 dengan *feature 1* (atribut nilai baca kitab), *feature 2* (atribut nilai nahwu raport), *feature 3* (atribut nilai b.arab raport) dengan data uji (plot bulat berwarna merah) dan data latih (plot x berwarna hijau muda). Hasil ilustrasi adalah gambar garis solid miring dengan garis pemisah nonlinier yang berjarak dekat.

Berdasarkan hasil ilustrasi, dapat diambil kesimpulan bahwa semakin kecil nilai *degree* maka semakin kecil kemungkinan terjadi kesalahan klasifikasi, dan semakin besar nilai *degree* maka semakin banyak kemungkinan terjadi kesalahan klasifikasi.

3.2.6 Pengaruh parameter C dan Gamma pada kernel Sigmoid

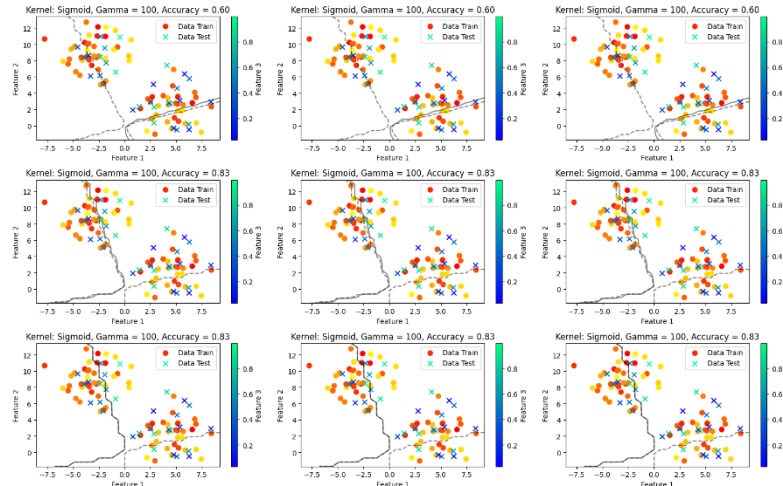
Berikut perbandingan nilai parameter C dengan nilai *Gamma* pada kernel Sigmoid berdasarkan akurasi yang dihasilkan:



Gambar 12. Pengaruh C dan Gamma pada kernel Sigmoid

Grafik pada **gambar 12** menunjukkan bahwa semakin besar nilai parameter *Gamma* yang digunakan pada kernel *Sigmoid* belum tentu menghasilkan akurasi terbaik dalam sebuah pemodelan, dan juga menghasilkan nilai akurasi yang cenderung sama dan konstan. Nilai

parameter *Gamma* pada penelitian ini paling optimal pada $Gamma=0,001$. Akurasi yang dihasilkan pada nilai $C=1$ dan $Gamma=0,001$ yaitu 71% pada penelitian ini. Nilai *Gamma* dan C default yang digunakan sendiri $Gamma=1$ dan $C=1$.



Gambar 13. Ilustrasi pada Kernel Sigmoid

Pada **gambar 13** adalah ilustrasi pada kernel *Sigmoid* dengan parameter C dan *Gamma*. Dalam percobaan ini model terbaiknya yaitu $C=1$ dan $Gamma=1$ dengan *feature 1* (atribut nilai baca kitab), *feature 2* (atribut nilai nahwu raport), *feature 3* (atribut nilai b.arab raport) dengan data uji (plot bulat berwarna merah) dan data latih (plot x berwarna hijau muda). Hasil ilustrasi adalah gambar garis pemisah non-linear yang membentuk pola acak pada plot bulat merah dan juga garis solid yang berada pada plot bulat kuning.

Berdasarkan hasil ilustrasi, dapat diambil kesimpulan bahwa semakin kecil nilai C dan *Gamma* maka semakin kecil kemungkinan terjadi kesalahan klasifikasi, dan semakin besar nilai C dan *Gamma* maka semakin banyak kemungkinan terjadi kesalahan klasifikasi.

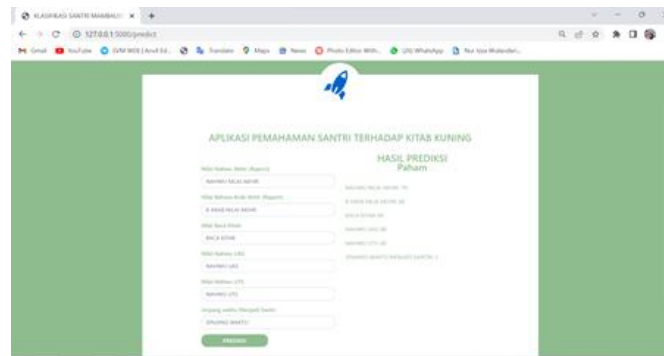
3. 4 Hasil Implementasi Web

Setelah melakukan uji coba sistem dan memilih model terbaik, maka selanjutnya melakukan implementasi pada web. Berikut adalah beberapa fitur yang terdapat pada implementasi web sistem ini:

3.4.1 Tampilan Awal Web, Input Nilai dan hasil prediksi



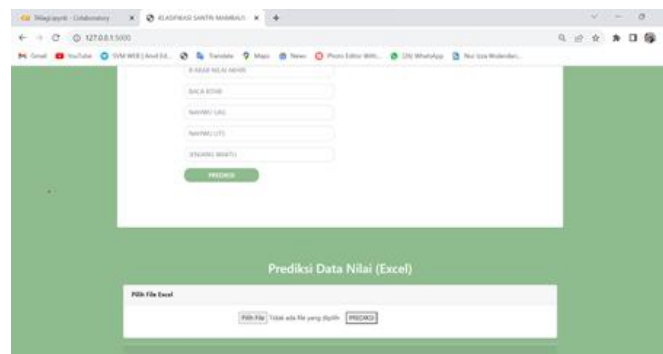
Gambar 14. Tampilan Input Nilai Website Pemahaman Santri



Gambar 15. Hasil Prediksi Paham

Pada **gambar 14** adalah tampilan awal saat mengunjungi Website Pemahaman Santri. Terdapat fitur inputan sejumlah 6, yaitu input nilai raport nahwu, raport bahasa arab, baca kitab, UAS dan UTS nahwu, dan jenjang waktu selama menjadi santri. Kemudian pada **gambar 17** adalah hasil prediksi paham.

3.4.2 Tampilan unggah data excel dan hasil prediksi



Gambar 16. Tampilan Input data excel Website Pemahaman Santri

	NAHWU NILAI ADHW	B. ARAB NILAI ADHW	BACA KITAS	NAHWU UAS	NAHWU UTS	JENJANG WAKTU	PREDIKSI
1	74	79	67	62	58	3	Paham
2	70	69	56	48	42	6	Tidak Paham
3	75	72	75	76	65	3	Paham
4	74	67	75	76	65	6	Paham
5	66	67	76	68	66	6	Tidak Paham
6	66	62	67	64	65	3	Paham
7	66	65	64	72	72	6	Paham
8	79	65	69	65	65	3	Paham
9	60	47	59	60	64	3	Tidak Paham

Gambar 17. Tampilan Hasil Prediksi dengan data excel

Pada **gambar 18** adalah tampilan saat akan mengunggah data excel nilai santri yang akan diprediksi. Setelah mengisi inputan tersebut pada bagian samping kanan terdapat tombol prediksi yang berfungsi memproses prediksi dari data nilai santri yang telah dimasukkan. Pada **gambar 19** adalah tampilan hasil prediksi dengan data excel yang telah diupload.

4 Kesimpulan dan Saran

Berdasarkan hasil analisis dari penelitian yang telah dilakukan dapat ditarik kesimpulan bahwa Metode *Support Vector Machine* dapat digunakan untuk Klasifikasi terhadap pemahaman dalam pembelajaran kitab kuning pada santri ponpes Mambaus Sholihin. Dari percobaan yang telah dilakukan dapat disimpulkan bahwa kernel *RBF* menjadi model terbaik dengan akurasi sebesar 89%, *presisi* sebesar 93% dan *recall* sebesar 83%. Pada kernel *RBF* hasil ilustrasi berupa gambar garis solid miring dengan garis pemisah nonlinear yang berjarak.

Saran yang dapat digunakan untuk pengembangan penelitian selanjutnya, yaitu dengan menambahkan atribut data pada nilai mata pelajaran bahasa arab agar klasifikasi lebih akurat dan menghasilkan akurasi lebih tinggi. Menambahkan nilai parameter pada keempat *hyperparameter*.

Daftar Pustaka

- [1] A. A. Mahfudh dan H. Mustofa, "Klasifikasi Pemahaman Santri Dalam Pembelajaran Kitab Kuning Menggunakan Algoritma Naive Bayes Berbasis Forward Selection," *Walisongo J. Inf. Technol.*, vol. 1, no. 2, hal. 101, 2019, doi: 10.21580/wjit.2019.1.2.4529.
- [2] A. Grosir, P. Toko, dan R. Bungursari, "Jurnal manajemen informatika," vol. 4, no. 2, hal. 18–24, 2017.
- [3] A. H. As dan N. Octavia, "Pembelajaran Kitab Kuning Menggunakan Algoritma C4 . 5 . Analisis Pohon Keputusan (Decison Tree) Di Pesantren," vol. 2, no. 3, hal. 266–269, 2021.
- [4] A. Perdana, M. Tanzil Furqon, dan Indiriati, "Penerapan Algoritma Support Vector Machine (SVM) Pada Pengklasifikasian Penyakit Kejiwaan Skizofrenia (Studi Kasus: RSJ. Radjiman Wediodiningrat, Lawang)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 9, hal. 3162–3167, 2018, [Daring]. Tersedia pada: <http://j-ptiik.ub.ac.id>
- [5] F. Hilmiyah, "Prediksi Kinerja Mahasiswa Menggunakan Support Vector Machine untuk Pengelola Program Studi di Perguruan Tinggi (Studi Kasus: Program Studi Magister Statistika ITS)," *Dep. Manaj. Teknol. Bid. Keahlian Manaj. Teknol. Inf. Fak. Bisnis Dan Manaj. Teknol. Inst. Teknol. Sepuluh Nop. Surabaya*, hal. 1–99, 2017, [Daring]. Tersedia pada: <http://repository.its.ac.id/46712/>
- [6] J. P. Maulana dan I. Irhamah, "Klasifikasi Kabupaten di Provinsi Jawa Timur Berdasarkan Indikator Daerah Tertinggal dengan metode Support Vector Machine (SVM) dan Entropy Based Fuzzy Support Vector Machine (EFSVM)," *Inferensi*, vol. 1, no. 1, hal. 9, 2018, doi: 10.12962/j27213862.v1i1.6715.
- [7] F. T. Informasi, "Online Shop Support Vector Machine (Studi Kasus : Bento Shop) Request for Online Shop Using Support Vector Machine Algorithm (Case Study : Bento Shop)".
- [8] I. S. Al-Mejibli, J. K. Alwan, dan D. H. Abd, "The effect of gamma value on support vector machine performance with different kernels," *Int. J. Electr. Comput. Eng.*, vol. 10, no. 5, hal. 5497–5506, 2020, doi: 10.11591/IJECE.V10I5.PP5497-5506.

- [9] H. I. Rhys, *Hefin I. Rhys*. 2020 "Machine Learning with R, the tidyverse and mlr".
- [10] M. Fajri dan A. Primajaya, "Komparasi Teknik Hyperparameter Optimization pada SVM untuk Permasalahan Klasifikasi dengan Menggunakan Grid Search dan Random Search," *J. Appl. Informatics Comput.*, vol. 7, no. 1, hal. 2548–6861, 2023, [Daring]. Tersedia pada: <http://jurnal.polibatam.ac.id/index.php/JAIC>

Dapat diakses online pada <http://jurnal.upnyk.ac.id/index.php/telematika/>
Telah terindeks oleh:



Home > User > Author > Active Submissions

Active Submissions

[ACTIVE](#) [ARCHIVE](#)

ID	MM-DD SUBMIT	SEC	AUTHORS	TITLE	STATUS
10685	08-25	GEN	Amalia	KLASIFIKASI PEMAHAMAN SWITTRI DALAM PENBELAJARAN KITAB...	Awaiting assignment

Start a New Submission
CLICK HERE to go to step one of the five-step submission process.

Refbacks

[ALL](#) [NEW](#) [PUBLISHED](#) [IGNORED](#)

DATE ADDED	HITS	URL	ARTICLE	TITLE	STATUS	ACTION
There are currently no refbacks.						

[Print](#) [Ignore](#) [Delete](#) [Select All](#)



USER

You are logged in as...
nabila23

- My Journals
- My Profile
- Log Out

NOTIFICATIONS

- View
- Manage

AUTHOR

Submissions

- Active (1)
- Archive (0)
- New Submission