

## Klasifikasi Penyakit Gigi Dan Mulut Menggunakan Metode Support Vector Machine

Ana Mariyam Puspitasari<sup>1</sup>, Dian Eka Ratnawati<sup>2</sup>, Agus Wahyu Widodo<sup>3</sup>

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya  
Email: <sup>1</sup>anamariyam6@gmail.com, <sup>2</sup>dian\_ilkom@ub.ac.id, <sup>3</sup>a\_wahyu\_w@gmail.com

### Abstrak

Penyakit gigi dan mulut merupakan salah satu penyakit yang berdampak serius bagi kesehatan manusia secara umum, karena gigi dan mulut merupakan tempat masuknya suatu kuman dan bakteri. Penanganan penyakit gigi dan mulut diharuskan untuk segera ditangani lebih cepat dan benar, namun tidak semua tim ahli gigi dapat dengan cepat melakukan penanganan dikarenakan kurangnya tim ahli gigi yang berada ditempat kerja atau rumah sakit selama 24 jam. Mengetahui jenis penyakit gigi dan mulut sejak awal sangatlah penting. Oleh karena itu diperlukan adanya sistem yang mempunyai kemampuan untuk mengklasifikasikan jenis penyakit gigi dan mulut guna membantu masyarakat dalam melakukan diagnosa awal terhadap penyakit gigi dan mulut. Dalam penelitian ini sistem klasifikasi yang digunakan yakni menggunakan metode SVM, karena metode SVM dapat mengatasi masalah klasifikasi dan regresi dengan linear maupun *non-linear* sehingga dapat menjadi suatu kemampuan algoritma pembelajaran pada klasifikasi ataupun regresi. Pada penelitian ini strategi yang digunakan yakni *One-Against-All* dan karena proses yang nantinya akan dilakukan bersifat *non-linear* sehingga kernel yang digunakan yakni kernel RBF. Hasil klasifikasi yang diperoleh dengan menggunakan metode SVM mempunyai rata – rata nilai akurasi sebesar 94.442% dengan menggunakan dataset sebanyak 122 data dan dengan parameter *sequential training SVM* nilai  $\lambda$  (*lamda*) = 0.1,  $\gamma$  (*gamma*) = 0.1,  $C$  (*Complexity*) = 1,  $\epsilon$  (*epsilon*) =  $1.10^{-10}$  dengan *itermax* = 50 dan rasio data 80%:20%. Dengan pencapaian hasil akurasi yang baik, maka penelitian ini dapat diterapkan untuk membantu melakukan klasifikasi penyakit gigi dan mulut dengan metode *support vector machine*.

**Kata kunci:** Klasifikasi, *Support Vector Machine*, *Radial Basis Function*, Penyakit Gigi dan Mulut.

### Abstract

Oral diseases is one of the most serious diseases that impact to human health in general, as the mouth is a place where the germ and bacteria oral diseases should be handled immediately but not all dental expert can quickly do the handling due to the lack of a dental expert that is available in the hospital for 24 hours. Knowing the types oral diseases since the beginning is very important. Therefore, a system that has the ability to classify types of oral diseases will be very helpful in order to help the community in conducting early diagnosis of oral diseases. This research used classification system using of SVM method because SVM method can resolve the problem of classification and regression with linear or *non-linear* kernel with its capability as a learning algorithm on the classification or regression. This research used *One-Against-All* strategies for *non-linear* process and used RBF kernel. The results obtained using SVM method has a mean median values of accuracy – 94,442% using the dataset as much as 122 data and with the parameter  $\lambda$  value SVM training sequential (*lamda*) = 0.1,  $\gamma$  (*gamma*) = 0.1,  $C$  (*Complexity*) = 1,  $\epsilon$  (*epsilon*) =  $1.10^{-10}$  with *itermax* = 50 and ratio data 80%: 20%. The results shows good accuracy, and the research can be applied to help perform classification of oral disease using support vector machine method.

**Keywords:** *Classification, Support Vector Machine, Radial Base Function, Oral diseases*

### 1. PENDAHULUAN

Kesehatan gigi dan mulut terkadang memang merupakan prioritas kesekian bagi

beberapa orang, padahal sebenarnya penyakit gigi dan mulut berdampak serius bagi kesehatan secara umum, sebab gigi dan mulut merupakan

tempat masuknya kuman dan bakteri sehingga kemungkinan besar dapat mengganggu kesehatan organ tubuh yang lainnya (Ratih, 2012).

Menurut Riskesdas tahun 2007 dan 2013 persentase penduduk Indonesia yang mengalami masalah pada gigi dan mulut meningkat dari 23.2% menjadi 25.9% dan yang menerima perawatan medis hanya sebesar 31.1% (Riskesdas, 2013).. Beberapa faktor yang menyebabkan timbulnya penyakit gigi dan mulut antara lain yakni mengkonsumsi rokok yang berlebihan dan kurang sehat sehingga membahayakan kesehatan, baik kesehatan gigi dan mulut maupun organ yang lain, pemakaian tembakau dan alkohol yang berlebihan sangat berbahaya, kurangnya menjaga kebersihan mulut, adanya jamur, adanya bakteri dan virus HIV (Ratih, 2012).

Penanganan terhadap penyakit gigi dan mulut diharuskan untuk dilakukan dengan cepat dan benar. Sebab dengan kita mengetahui jenis penyakit gigi dan mulut sejak awal sangatlah Penting, guna melakukan proses penyembuhan dengan lebih cepat dan tepat. Oleh karena itu perlu adanya sistem cerdas yang mampu mengklasifikasikan penyakit gigi dan mulut berdasarkan gejalanya. Dengan adanya sistem cerdas diharapkan dapat membantu pengguna maupun tim medis untuk mengetahui jenis penyakit gigi dan mulut dan dapat menentukan langkah awal untuk menangani penyakit tersebut. Meskipun seorang dokter gigi dan mulut adalah seorang yang ahli dibidangnya, namun terkadang manusia biasan juga mempunyai keterbatasan daya ingat dan stamina kerja. Sehingga yang ditakuti yakni ketika seorang dokter gigi dan mulut mungkin saja melakukan kesalahan pada saat mengambil hasil diagnosa yang dapat berakibat fatal. Beberapa penelitian telah dilakukan dengan memilih metode klasifikasi terbaik, yakni dilakukan oleh Akbar Afizal Laksita pada tahun 2015 terhadap penyakit stroke dengan menggunakan metode *Support Vector Machine*. Penelitian tersebut bertujuan untuk membuat hyperplane yang optimal, sehingga diperoleh hasil rata – rata akurasi sebesar 0.8939 pada saat panjang fold 5 dengan nilai parameter epsilon 0.001 dan cost 10. Kemudian penelitian yang kedua yakni dilakukan oleh A.Muis, dkk., pada tahun 2015 mengenai klasifikasi tweet dengan

menggunakan kernel *Radial Basis Function (RBF)*. Hasil yang diperoleh pada penelitian tersebut adalah nilai akurasi tertinggi 97.54% untuk data yang belum dilakukan pemilihan fitur, sedangkan yang sudah dilakukan pemilihan fitur mencapai nilai akurasi tertinggi 99.12%. Dari beberapa penjelasan masalah dan penelitian yang telah dilakukan diatas, penulis ingin melakukan penelitian terhadap penyakit gigi dan mulut dengan menggunakan metode *support vector machine*. Diharapkan penelitian ini dapat menghasilkan suatu sistem cerdas yang dapat menghasilkan kesimpulan apakah seseorang mempunyai potensi terkena penyakit gigi dan mulut jenis yang mana dengan menggunakan metode *support vector machine*.

## 2. GIGI DAN MULUT

### 2.1 Rongga Mulut

Didalam rongga mulut terdapat beberapa bagian seperti gigi, lidah, gusi, bibir dan jaringan lunak lainnya yang mempunyai fungsi masing – masing.

- Bibir mempunyai fungsi dapat menjaga makanan dan minuman agar tidak berceceran keluar mulut, membantu kita dalam berbicara, dll.
- Gusi mempunyai fungsinya yaitu untuk melindungi serat – serat halus yang mengikat akar gigi kepada tulang rahang (Kesehatan, 2012)
- Lidah mempunyai fungsinya yakni sebagai alat indera perasa sehingga dapat merasakan makanan dan minuman.
- Gigi memiliki fungsi dari gigi yakni dapat mengunyah makanan, menghancurkan makanan, dll.
- Jaringan lunak merupakan jaringan yang meliputi bagian pipi, bibir, langit – langit dan jaringan lunak yang lainnya dibawah lidah

#### 2.1.2 Gigi

Bagian dari gigi yakni mahkota gigi dan akar gigi yang mempunyai banyak fungsi. Macam – macam bentuk gigi beserta fungsinya antara lain:

- Gigi seri yang berfungsi untuk memotong makanan yang akan masuk ke dalam mulut.
- Gigi taring mempunyai fungsi mencabik – cabik makanan sesudah dipotong

- c. Gigi geraham mempunyai fungsi menggiling atau menghaluskan makanan.

## 2.2 Penyakit Gigi dan Mulut

Gigi dan mulut merupakan bagian dari tubuh kita yang sangat vital, sebab disanalah tempat masuknya makanan yang kita makan dan gigi yang menghancurkan makanan tersebut. Oleh sebab itu kesehatan dan kebersihan gigi dan mulut sangatlah penting. Banyak faktor yang dapat menyebabkan timbulnya penyakit gigi dan mulut, antara lain seperti diet yang tidak sehat, mengkonsumsi minuman alkohol dan merokok yang berbahaya dan berlebihan, dan kebersihan mulut yang tidak terawat, jamur dan bakteri. Beberapa macam penyakit gigi dan mulut yang biasa dijumpai antara lain :

- a. Gingivitis merupakan penyakit radang gusi yang mengalami pembengkakan pada mulut sebab kurang terjaganya kebersihan mulut sehingga menyebabkan adanya karang – karang gigi atau plak yang menumpuk dan berbatasan dengan tepi gusi (Lita, 2016).
- b. *Acute Necrotizing Ulcerative Gingivitis (ANUG)* adalah penyakit yang disebabkan oleh adanya infeksi pada nekrosis gingiva. Penyakit ini dapat terjadi pada siapa saja, terutama orang yang mengkonsumsi rokok secara berlebihan, stress berat, dan malnutrisi berat, dll
- c. Karies gigi merupakan penyakit gigi yang terjadi pada kerusakan jaringan gigi hingga membentuk lubang
- d. Pulpitis merupakan proses radang pada jaringan pulpa gigi yang menetap, gejalanya yakni gigi nyeri ketika mendapat rangsangan panas atau dingin
- e. Nekrosis Pulpa adalah penyakit gigi yang disebabkan oleh adanya bakteri, trauma dan iritasi yang menyebabkan kerusakan dan kematian pada pulpa yang disebabkan oleh pulpitis yang tidak dirawat (Yamin, 2012)
- f. Periodontitis merupakan inflamasi jaringan dan infeksi yang terjadi pada gingiva (gingivitis) yang tidak dirawat dan menyebar ke ligamen dan tulang alveolar penyangga gigi.
- g. Herpes Simpleks adalah infeksi virus HIV yang terjadi pada sudut bibir atau mulut. Gejala yang ditimbulkan antara lain

sensitive, terbakar pada daerah bibir atau perbatasan kulit bibir

- h. Glositis merupakan penyakit radang pada lidah dimana keadaannya di dalam mulut biasanya ditunjukkan dengan adanya pembengkakan di lidah, jika kasusnya lebih parah mampu memicu penyumbatan pernafasan pada saat lidah membengkak yang sangat parah (Lita, 2016).
- i. Impaksi gigi adalah kerusakan erupsi pada gigi yang disebabkan adanya malposisi, kekurangan tempat atau terhalangi gigi yang lain. Hal itu disebabkan oleh adanya gusi bengkak, demam, dan gigi yang tumbuh tidak sempurna

## 2.3 Klasifikasi

Klasifikasi merupakan suatu proses yang bertujuan untuk menentukan suatu obyek kedalam suatu kelas atau kategori yang sudah ditentukan sebelumnya. Menurut (Elly Susilowati, 2015) klasifikasi adalah proses dari pembangunan terhadap suatu model yang mengklasifikasikan suatu objek sesuai dengan atribut – atributnya. Klasifikasi data ataupun dokumen juga dapat dimulai dari membangun aturan klasifikasi tertentu yang menggunakan data training yang sering disebut sebagai tahapan pembelajaran dan pengujian digunakan sebagai data testing. (Winarko2, Oktober 2014).

Beberapa tugas dari klasifikasi yang melibatkan proses pembangunan terhadap model yang dibentuk untuk melakukan prediksi target atau variabel dari data set yang sudah jelas, ataupun variabel independen. Klasifikasi juga dapat dilakukan dengan menggunakan beberapa metode atau berbagai jenis pengklasifikasian. Beberapa metode yang sering digunakan pada klasifikasi yakni *decision-tree*, *rule based*, *ANN*, *nearest-neighbor*, dan *naive Bayesian* (Weiss, 2010).

## 2.4 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. SVM pertama kali diperkenalkan pada tahun 1992 oleh Vapnik sebagai rangkaian dari beberapa konsep – konsep unggulan dalam bidang pattern recognition (Elly Susilowati,

2015)

Tingkat akurasi pada model yang akan dihasilkan oleh proses peralihan dengan SVM sangat bergantung terhadap fungsi kernel dan parameter yang digunakan (Siagian, 2011) 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 yang berdimensi tinggi (F, 2012)

**2.5 Kernel Trick**

Pada umumnya untuk masalah yang ada dalam domain dunia nyata, kebanyakan bersifat non linier. Algoritma *Support Vector Machine (SVM)* dimodifikasi dengan cara memasukkan fungsi kernel kedalam non linier SVM, dengan cara yang pertama yaitu data  $x_i$  dipetakan ke dalam fungsi  $\Phi(x)$  ke ruang vector yang memiliki ukuran dimensi tinggi. Notasi matematika dari mapping akan ditunjukkan seperti pada rumus berikut (Nugroho, 2003)

$$\Phi : \mathbb{R}^d \rightarrow \mathbb{R}^q \quad d < q \quad (1)$$

Dikarenakan transformasi  $\Phi$  pada umumnya tidak diketahui, oleh karena itu fungsi dari Kernel Trick dapat digantikan sesuai rumus berikut (Nugroho, 2003):

$$\kappa(\vec{x}_i, \vec{x}_j) = \Phi(\vec{x}_i) \cdot \Phi(\vec{x}_j) \quad (2)$$

Fungsi dari nilai  $\Phi(\vec{x}_i)$  pada dot *product* dengan menggunakan dua vector yang menunjukkan *feature* dari atribut sehingga dapat di hitung dengan baik pada *feature space*. Selanjutnya *feature space* akan dibuat sebuah fungsi linear yang mewakili fungsi dari non-linear pada *input space*. Di dalam *input space* tidak bisa dipisahkan secara linear, namun *input space* bisa dipisahkan di *feature space* dan dapat membantu proses klasifikasi menjadi lebih mudah. Untuk mendapatkan solusi pada fungsi klasifikasi dari data (x), didapatkan rumus berikut (Nugroho, 2003):

$$f(x) = \sum_{i=1, \vec{x}_i \in SV}^n a_i y_i \kappa(\vec{x}, \vec{x}_i) + b \quad (3)$$

dimana

$$\begin{aligned} \vec{w} &= \sum_{i=1, \vec{x}_i \in SV}^n a_i y_i \Phi(\vec{x}) \cdot \Phi(\vec{x}_i) \\ &= \sum_{i=1, \vec{x}_i \in SV}^n a_i y_i \kappa(\vec{x}, \vec{x}_i) \end{aligned} \quad (4)$$

Maksud dari persamaan diatas yakni dengan subset dari data training dot yang terpilih sebagai *support vector*, dengan kata lain data () yang

berkorespondensi terhadap  $\alpha \geq 0$ . Jenis kernel yang umum digunakan seperti yang terlihat pada Tabel 1 :

**Tabel 1 Kernel yang sering digunakan**

Jenis Kernel	Definisi
Polynomial	$\kappa(\vec{x}_i, \vec{x}_j) = (\vec{x}_i \cdot \vec{x}_j + 1)^p$
RBF	$\kappa(\vec{x}_i, \vec{x}_j) = \exp(-\frac{\ \vec{x}_i - \vec{x}_j\ ^2}{2\sigma})$
Sigmoid	$\kappa(\vec{x}_i, \vec{x}_j) = \tanh(\alpha \vec{x}_i \cdot \vec{x}_j + \beta)$

Penggunaan kernel dapat dilakukan dalam percobaan untuk menentukan parameter kernel dan menghasilkan keakuratan yang terbaik dalam proses klasifikasi. Kernel linear digunakan pada saat data yang diklasifikasikan dapat dengan mudah dipisahkan dengan sebuah garis atau *hyperplane*, sementara untuk kernel *non-linear* digunakan pada saat data yang digunakan dipisah dengan menggunakan garis lengkung atau sebuah bidang pada ruang yang mempunyai dimensi tinggi. Dalam penelitian ini *kernel* yang digunakan yakni *kernel* RBF, dimana kernel RBF merupakan kernel yang mempunyai performansi yang baik pada parameter tertentu dengan kesalahan pelatihan yang minimum.

**2.6 Sequential Training**

Salah satu Hyperplane yang optimal terletak didalam SVM yang dapat ditemukan dengan merumuskan ke dalam *Quadratic Programming (QP)* problem yang nantinya akan diselesaikan dengan analisa numerik. Alternatif lain yang cukup simple dan sederhana yakni *Sequential Training* seperti halnya yang telah dikembangkan oleh (Vijayakumar S, 1999) sebagai berikut:

- a. Inisialisasi  $\alpha_1 = 0$  setelah itu hitung matrik *Hessian*. Matrik *Hessian* adalah perkalian antara kernel gaussian dengan nilai Y. Nilai Y disini yaitu nilai berupa vector yang berisi nilai 1 dan -1.  $\alpha_1$  digunakan untuk mencari nilai *support vector*. Untuk setiap data dari i sampai j, hitung menggunakan rumus Matrik *Hessian* yang ditunjukkan seperti berikut:

$$D_{ij} = y_i y_j (K(x_i, x_j) + \lambda^2) \quad (5)$$

- b. Kemudian melakukan rumus dibawah ini:

$$E_i = \sum_{j=1}^i a_j D_{ij} \quad (6)$$

$$\begin{aligned} \delta \alpha_i &= \min \{ \max[y(1 - E_i) - \alpha_i] \\ &C - \alpha_i \} \end{aligned} \quad (7)$$

$$\alpha_i = \alpha_i + \delta\alpha_i \tag{8}$$

Keterangan:

$\alpha_i$  = alfa ke- $i$

$D_{ij}$  = matriks Hessian

$E_i$  = error rate

$C$  = konstanta  $C$

$\delta\alpha_i$  = delta alfa ke- $i$

Setelah itu kembali ke langkah 2 sampai nilai  $\alpha_i$  mencapai konvergen. Konvergensi didapatkan dari tingkat perubahan pada nilai  $\alpha_i$ .

### 2.7. One Against All

Strategi dari menggunakan metode *One-Against-All* ini yaitu membangun sejumlah nilai  $k$  kedalam model SVM biner ( $k$  merupakan jumlah kelas). Dari setiap klasifikasi ke- $i$  dilatih dengan menggunakan data secara keseluruhan. Misal, terdapat sebuah permasalahan klasifikasi dengan 4 buah kelas yang digunakan pada pelatihan hanya 4 buah SVM biner : (Sembiring, September 2007) seperti yang ditunjukkan pada tabel berikut :

**Tabel 2 Contoh 4 SVM dengan metode *One-Against-All***

$Y_i = 1$	$Y_i = -1$	Hipotesis
Kelas 1	Bukan kelas 1	$f^1(x) = (w^1)x + b^1$
Kelas 2	Bukan kelas 2	$f^2(x) = (w^2)x + b^2$
Kelas 3	Bukan kelas 3	$f^3(x) = (w^3)x + b^3$
Kelas 4	Bukan kelas 4	$f^4(x) = (w^4)x + b^4$

### 2.8 Kelebihan dan Kekurangan SVM

Beberapa kelebihan yang ada pada SVM antara lain:

- Generalisasi yang artinya mempunyai kemampuan suatu metode (SVM, neural network, dsb) untuk melakukan klasifikasi suatu pattern, dimana tidak termasuk data yang dipakai dalam fase pembelajaran metode tersebut.
- Curse of dimensionality* yakni suatu masalah yang biasa dihadapi suatu metode pattern recognition dalam melakukan estimasi parameter. Dikarenakan jumlah sampel data yang relatif sedikit jika dibandingkan dengan dimensional ruang vector data, sehingga semakin tinggi dimensi ruang vector yang diolah, maka akan menimbulkan konsekuensi yang

membutuhkan jumlah data dalam berdimensi tinggi juga.

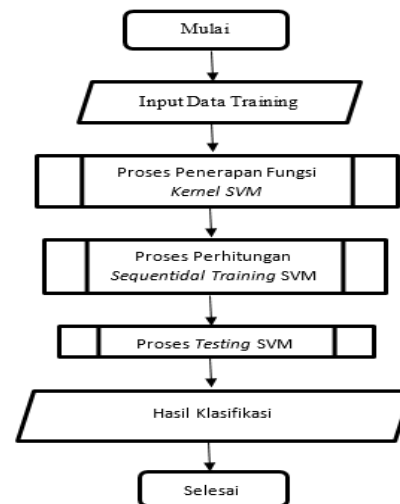
- Feasibility*. Metode SVM dengan mudah dapat diimplementasikan karena proses dalam menentukan support vector dapat dirumuskan dalam QP problem.

Beberapa kelemahan pada SVM antara lain:

- Sulit digunakan pada problem yang mempunyai ukuran besar. Dimaksudkan ukuran besar dengan jumlah sample yang diolah
- Metode SVM secara teoritik dikembangkan dengan fungsi untuk problem klasifikasi yang menggunakan dua class.

### 3. PERANCANGAN DAN IMPLEMENTASI

Proses algoritma SVM yang harus dilakukan pertama kali yaitu mengambil dataset. Kemudian dilanjutkan dengan proses perhitungan kernel SVM, proses training SVM, proses testing SVM, dan yang terakhir yaitu tahap evaluasi klasifikasi. Gambar 1 dipaparkan alur proses algoritma *Support Vector Machine*.



Gambar 1. Alur Proses Algoritma Support Vector Machine

Berdasarkan Gambar 1 proses algoritma *support vector machine* tahapan pertama yakni menginputkan data training format .xls. Data training di bagi secara acak dari setiap kelas sehingga minimal terdapat 3 data yang mewakili data training dan data testing dari setiap kelas. Setelah menginputkan data set, langkah selanjutnya yakni melakukan perhitungan kernel. Disini kernel yang digunakan yaitu kernel RBF. Setelah melakukan proses perhitungan kernel, langkah selanjutnya yaitu

melakukan perhitungan data training dengan menggunakan *Sequential Training SVM*. Setelah melakukan tahapan perhitungan *sequential training SVM*, langkah selanjutnya yaitu melakukan perhitungan testing SVM, dan akan menghasilkan keluaran berupa hasil klasifikasi.

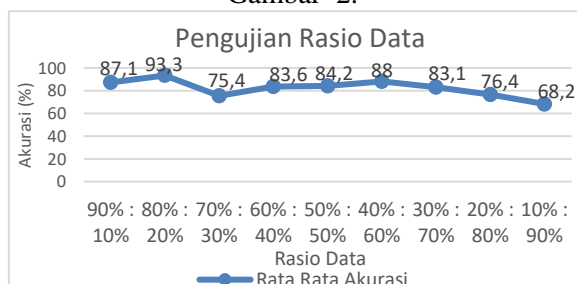
#### 4. PENGUJIAN DAN ANALISIS

Pada penelitian ini dilakukan pengujian terhadap rasio data, pengujian terhadap *kernel*, pengujian parameter *lamda*, pengujian parameter *gamma*, pengujian parameter *complexity*, dan pengujian terhadap jumlah iterasi.

##### 4.1 Pengujian Rasio Data

Pada pengujian pertama digunakan untuk mengetahui perbandingan tingkat akurasi dari dataset yang digunakan. Dataset yang akan digunakan baik data training maupun data testing yang berjumlah sebanyak 122 dan akan dibagi sesuai dengan perbandingan rasio yang telah ditentukan guna untuk melihat hasil akurasi terbaik. Perbandingan rasio yang akan digunakan yaitu rasio 90%:10%, 80%:20%, 70%:30%, 60%:40%, 50%:50%, 40%:60%, 30%:70%, 20%:80%, 10%:90%. Pengujian juga dilakukan dengan menggunakan *kernel RBF* dan nilai parameter pada *sequential training* yang digunakan dalam pengujian ini yaitu nilai  $\lambda = 0.5$   $\gamma = 0.01$   $C = 1$   $\epsilon = 1.10^{-10}$  iterasi = 50

Gambar 2.



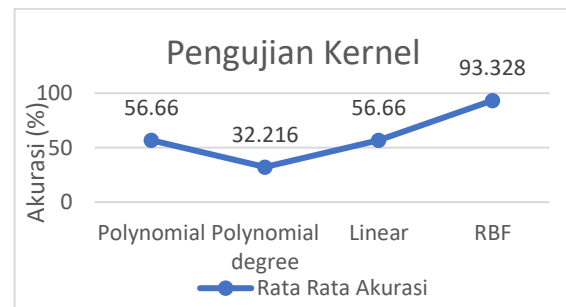
Grafik 2. Hasil Pengujian Rasio Data

Berdasarkan Grafik 2 diperoleh nilai akurasi yang terbaik terdapat pada rasio data 80%:20% dengan nilai akurasi sebesar 93.328%. Data training dan data testing yang digunakan dipilih secara acak dari setiap kelas, sehingga minimal terdapat tiga data yang mewakili data training dan data latih dari setiap kelas untuk data uji yang digunakan. Sehingga dapat disimpulkan bahwa secara keseluruhan algoritma *support vector machine* menghasilkan nilai tingkat akurasi yang lebih tinggi dan lebih

konsisten. Hal ini disebabkan oleh komposisi data yang digunakan pada data training maupun data testing tidak jauh berbeda.

##### 4.2 Pengujian Kernel

Pengujian terhadap jenis kernel yang digunakan, dimana pengujian tersebut dilakukan guna untuk mengetahui jenis kernel yang mana yang memperoleh hasil terbaik. Nilai parameter yang digunakan pada *sequential training SVM* pada pengujian ini adalah nilai  $\lambda = 0.5$   $\gamma = 0.01$   $C = 1$   $\epsilon = 1.10^{-10}$  iterasi = 50 dengan rasio data 80%:20%.

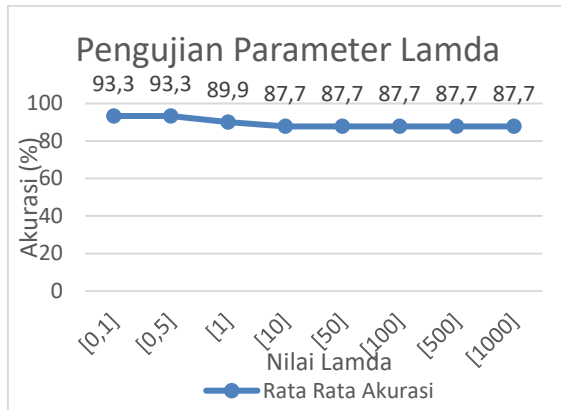


Grafik 3. Hasil Pengujian Kernel

Berdasarkan Grafik 3 diperoleh nilai akurasi yang paling optimal terdapat pada *kernel RBF* dengan nilai akurasi sebesar 93.328. Hal ini dapat disimpulkan bahwa *kernel RBF* lebih cocok digunakan pada data seperti jenis penelitian ini dibandingkan dengan jenis *kernel polynomial degree*, *polynomial degree 2*, dan *kernel linear*.

##### 4.3 Pengujian $\lambda$ (lamda)

Pengujian parameter lamda dilakukan guna untuk mengetahui skenario mana yang memperoleh hasil terbaik untuk nilai parameter lamda yakni 0.1, 0.5, 1, 10, 50, 200, 500, 1000. Nilai parameter yang digunakan pada *sequential training SVM* pada pengujian ini adalah nilai  $\gamma = 0.01$   $C = 1$   $\epsilon = 1.10^{-10}$  iterasi = 50 dengan rasio data 80%:20%.

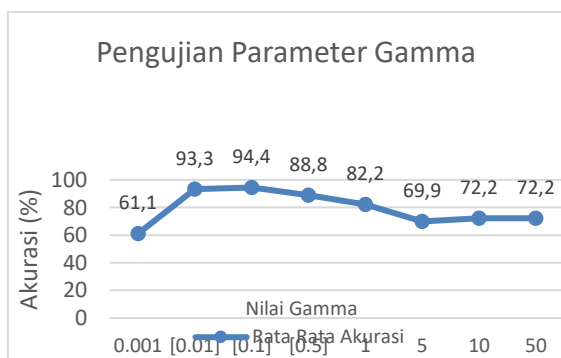


Grafik 4. Hasil Penguujian Parameter Lamda

Berdasarkan Grafik 4 diketahui bahwa nilai rata – rata tingkat akurasi yang paling tinggi sebesar 93.33 yaitu terdapat pada nilai lamda 0.1 dan 0.5. Disini dapat disimpulkan bahwa semakin besar nilai lamda tidak juga membuat nilai akurasi menjadi baik, karena apabila semakin besar nilai lamda akan membuat proses komputasi pada tahap perhitungan matriks hessian semakin lama. Hal ini disebabkan oleh adanya augmented factor (lamda) yang menjadikan sistem menjadi sangat lambat dalam mencapai nilai konvergen sehingga dapat menimbulkan ketidakstabilan dalam proses pembelajaran yang dilakukan.

#### 4.4 Penguujian $\gamma$ (Gamma)

Penguujian terhadap parameter sequential training SVM yaitu parameter  $\gamma$  (gamma). Penguujian tersebut dilakukan guna untuk mengetahui skenario mana yang memperoleh hasil terbaik untuk nilai parameter gamma yakni 0.001, 0.01, 0.1, 0.5, 1, 5, 10, 50. Nilai parameter yang digunakan pada sequential training SVM pada penguujian ini adalah nilai  $\lambda = 0.1$   $C = 1$   $\epsilon = 1.10^{-10}$  iterasi = 50 dengan rasio data 80%:20%.



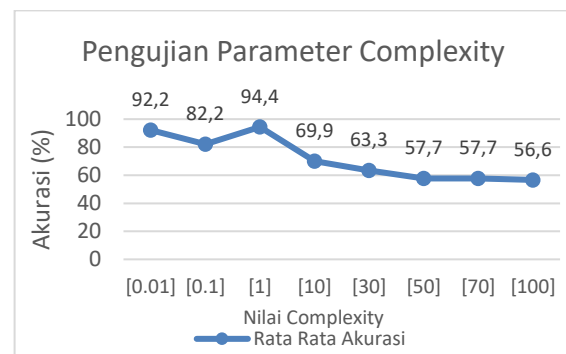
Grafik 5. Hasil Penguujian Parameter Gamma

Berdasarkan Grafik 5 diketahui bahwa rata

– rata nilai tingkat akurasi yang paling tinggi sebesar 94.442 yaitu terdapat pada nilai gamma 0.1 sedangkan untuk nilai akurasi terbaik sebesar 100 pada nilai gamma 0.01, 0.1, 0.5, dan 1. Pada penguujian parameter  $\gamma = 0.001$  mendapatkan hasil dengan nilai rata – rata akurasi yang sangat kecil, hal ini disebabkan nilai yang diperoleh dari perhitungan kernel RBF sangat besar sehingga menyebabkan algoritma menjadi sulit dalam mendapatkan kekonsistenan dari hyperplane

#### 4.5 Penguujian C (Complexity)

Penguujian terhadap parameter sequential training SVM yaitu parameter C (Complexity). Penguujian tersebut dilakukan guna untuk mengetahui skenario mana yang memperoleh hasil terbaik untuk nilai parameter complexity yakni 0.01, 0.1, 1, 10, 30, 50, 70, 100. Nilai parameter yang digunakan pada sequential training SVM pada penguujian ini adalah nilai  $\lambda = 0.1$   $\gamma = 0.1$   $\epsilon = 1.10^{-10}$  iterasi = 50 dengan rasio data 80%:20%



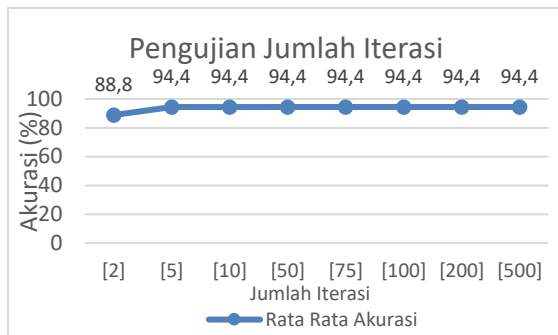
Grafik 6. Hasil Penguujian Parameter Complexity

Berdasarkan Grafik 6 pada Gambar 6 diketahui bahwa nilai rata – rata tingkat akurasi yang paling tinggi yaitu sebesar 94.442. Penguujian ini digunakan dengan tujuan untuk meminimalkan nilai error dan memperkecil nilai slack variabel, hal itu relatif penting diperlukan untuk memaksimalkan margin dan meminimalkan jumlah slack. Apabila nilai C mendekati angka 0 maka lebar margin yang terdapat pada bidang pemisah (hyperplane) menjadi maksimal, sehingga nilai C dapat digunakan untuk memperkecil nilai error pada proses training saat perhitungan nilai  $w$  (weight) dan nilai bias.

#### 4.6 Penguujian Jumlah Iterasi

Penguujian terhadap jumlah iterasi. Penguujian tersebut dilakukan guna untuk

mengetahui skenario mana yang memperoleh hasil terbaik. Nilai parameter yang digunakan pada *sequential training* SVM pada pengujian ini adalah nilai  $\lambda = 0.1$   $C = 1$   $\gamma = 0.1$   $\epsilon = 1.10^{-10}$  iterasi = 50 dengan rasio data 80%:20%.



Grafik 7. Hasil Pengujian Jumlah Iterasi

Berdasarkan Grafik 7 diketahui bahwa nilai rata – rata tingkat akurasi tinggi yaitu sebesar 94.442. Pada pengujian ini dengan data yang digunakan dapat disimpulkan bahwa terjadinya konvergen mulai pada iterasi ke 5. Hal itu dikarenakan saat jumlah iterasi bertambah terjadi, rasio *support vector* berjalan seimbang dan beberapa data tidak terletak jauh berbeda dari bidang pemisah (*hyperplane*). Namun pada pengujian ini nilai iterasi maksimum sangat berpengaruh terhadap perubahan nilai pada nilai  $\alpha$  (*alpha*) karena diperlukan untuk mendapatkan nilai  $\alpha_i$  yang konvergen.

## 5. KESIMPULAN

Berdasarkan hasil penelitian mengenai klasifikasi penyakit gigi dan mulut dengan menggunakan metode *support vector machine* dapat disimpulkan bahwa:

1. Metode *Support Vector Machine* dapat diterapkan pada permasalahan klasifikasi penyakit gigi dan mulut dengan hasil terbaik dan memberikan hasil akurasi sebesar rata rata 94.442% dengan nilai parameter pada *sequential training* SVM yaitu  $\lambda$  (*lamda*) = 0.1,  $\gamma$  (*gamma*) = 0.1,  $C$  (*Complexity*) = 1,  $\epsilon$  (*epsilon*) =  $1.10^{-10}$  dengan *itermax* = 50 dan rasio data 80%:20% dengan memperoleh hasil pengujian yang mulai konvergen pada jumlah iterasi ke 5.
2. Proses klasifikasi penyakit gigi dan mulut ini menggunakan *dataset* yang masih terbatas yakni sebesar 122 data untuk semua kelas sehingga terdapat sekitar 30 data untuk setiap kelas dengan jumlah parameter sebanyak 16. Pada penelitian ini

menggunakan empat kelas yaitu kelas pulpitis, gingivitis, nekrosis pulpa dan periodontitis.

3. Pada pengujian kernel hasil yang paling optimal terdapat pada kernel RBF dengan nilai akurasi rata – rata sebesar 93.329%. Dapat disimpulkan bahwa kernel RBF lebih cocok pada data jenis penelitian ini sebab kernel RBF mampu memprediksi kelas pada data itu sendiri, yang artinya mampu mengklasifikasikan tepat sesuai dengan kelas aslinya, dan lebih mempunyai performa lebih baik jika dibandingkan dengan kernel linear atau polynomial. Hal ini disebabkan karena data yang ada pada penelitian ini memiliki pola yang cenderung tetap, sedangkan kernel polynomial lebih cenderung pada data yang mempunyai sebaran data yang cenderung tidak tetap atau pola yang naik turun.

## 6. DAFTAR PUSTAKA

- Elly, S., Mira, K. S. & Alfian, A. G., 2015. *Implementasi Metode Support Vector Machine untuk Melakukan Klasifikasi Kemacetan Lalu Lintas Pada Twitter*. Makassar, S1 Universitas Telkom.
- F, R., 2012. *Perbandingan Klasifikais Tingkat Keganasan Breast Cancer dengan Menggunakan Regresi Logistik Ordinal dan Support Vector Machine*. s.l., s.n.
- Imelda, A. M. & Muhammad, A. M., Juni 2015. Penerapan Metode Support Vector Machine (SVM) Menggunakan Kernel Radial Basis Function (RBF) Pada Klasifikasi Tweet. *Jurnal Sains, Teknologi dan Industri*, Vol. 12, No. 2(ISSN 1693-2390), p. pp.189 – 197.
- Kesehatan, K., 2012. *Buku Panduan Pelatihan Kader Kesehatan Gigi dan Mulut di Masyarakat*. Republik Indonesia, s.n.
- Laksita, A. A., (2015). *Implementasi Algoritma SVM (Support Vector Machine) untuk Mengetahui Tingkat Resiko Penyakit Stroke*. Malang, S1 Universitas Brawijaya.
- Lita, 2016. *Penyakit Gigi dan Mulut*. [Online] Available at: <http://halosehat.com/penyakit/penyakit-gigi-dan-mulut/jenis-jenis-penyakit-gigi-dan-mulut> [Diakses 9 april 2017].



