

Penerapan Algoritma *Support Vector Machine* (SVM) Pada Pengklasifikasian Penyakit Kejiwaan Skizofrenia (Studi Kasus: RSJ. Radjiman Wediodiningrat, Lawang)

Arya Perdana¹, Muhammad Tanzil Furqon², Indriati³

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya

Email: ¹ aryaperd@gmail.com, ² m.tanzil.furqon@ub.ac.id, ³ indriati.tif@ub.ac.id

Abstrak

Skizofrenia merupakan penyakit yang menyerang kejiwaan seseorang, sehingga mengakibatkan tingkah laku dengan pola pikir tidak sesuai. Salah satu penyebab seseorang menderita skizofrenia adalah stress dan juga memiliki tekanan hidup yang berat dari berbagai aspek kehidupan. *Support Vector Machine* (SVM) adalah suatu algoritme yang dapat mengklasifikasikan jenis penyakit skizofrenia. Data yang digunakan pada penelitian ini adalah sebanyak 11 data yang terbagi menjadi 5 kelas. Kelas pada penelitian ini mewakili 5 jenis penyakit pada Skizofrenia yaitu paranoid, hebefrenik, katatonik, *undifferentiated*, dan simpleks. Pada dasarnya algoritme SVM merupakan metode klasifikasi linear, sehingga digunakan suatu *kernel* untuk mengatasi data yang bersifat nonlinear. Pada penelitian ini digunakan juga konsep *One Against All* untuk mengatasi permasalahan *multiclass*. Hasil akhir dari penelitian ini menghasilkan rata-rata akurasi tertinggi sebesar 50,09 %, dengan nilai konstanta $\lambda = 1$; $C = 0,1$; $\gamma = 0,1$; $itermax = 100$; $\varepsilon = 0,01$; dan juga menggunakan *kernel polynomial*. Pengujian pada penelitian ini menggunakan pengujian *K-Fold Cross Validation*, dengan menggunakan 11 *fold*.

Kata kunci: Skizofrenia, *Support Vector Machine*, *One Against All*.

Abstract

Schizophrenia is a disease that attacks a person's psyche, and resulting in behavior with an inappropriate mindset. One of the causes of a person suffering from schizophrenia is stress and also has severe life pressures from various aspects of life. Support Vector Machine (SVM) is an algorithm that can classify types of schizophrenia. The data used in this research is as much as 11 data which is divided into 5 classes. Classes in this study represent five types of diseases in schizophrenia are paranoid, hebefrenik, catatonic, undifferentiated, and simplex. Basically SVM algorithm is a method of linear classification, so that a kernel is used to overcome nonlinear data. In this research is also used One Against All concept to solve multiclass problem. The end result of this research resulted in the highest accuracy of 50.09%, with constant value $\lambda = 1$; $C = 0,1$; $\gamma = 0,1$; $itermax = 100$; $\varepsilon = 0.01$; and also uses polynomial kernels. Tests in this study using K-Fold Cross Validation test, using 11 fold.

Keywords: *Schizophrenia, Support Vector Machine, One Against All.*

1. PENDAHULUAN

Skizofrenia adalah salah satu penyakit gangguan jiwa. Penderita biasanya memiliki perbedaan antara pola pikir seseorang dengan tingkah lakunya (Maryatun, 2015). Pada umumnya penderita penyakit ini tidak dapat melakukan kegiatan sosial lagi dengan baik, hal ini dikarenakan tingkah laku penderita tidak lagi sesuai dengan tingkah laku manusia biasa pada hakikatnya. Di Indonesia sendiri masih banyak penderita penyakit skizofrenia yang belum dapat

diobati. Hal ini dikarenakan banyak masyarakat yang belum begitu paham dan mengerti tentang penyakit ini, sehingga masyarakat belum dapat mengetahui bagaimana cara mencegah dan mengatasi penyakit ini. Salah satu dampak negatif yang dapat ditimbulkan apabila pencegahan pada penyakit ini tidak dapat diatasi adalah bunuh diri. Hal ini dapat disebabkan dari stigma negatif yang diberikan oleh masyarakat terhadap penyandang penyakit ini.

Menurut Hendarsyah (2016), penyakit skizofrenia diklasifikasikan dalam beberapa

jenis, seperti skizofrenia *paranoid*, *hebefrenik*, *katatonik*, *undifferentiated*, dan *simpleks*. Penyakit ini diklasifikasikan berdasarkan gejala-gejala yang dijadikan parameter. Penyakit skizofrenia perlu diklasifikasikan, karena dalam menentukan jenis penyakit skizofrenia yang diderita pada pasien terbilang sulit. Alasan kenapa penyakit ini terbilang sulit dikarenakan banyaknya gejala pada penyakit ini, sehingga penanganannya juga dapat terhambat. Untuk menyelesaikan masalah tersebut maka dibutuhkan suatu sistem yang dapat membantu dalam mengklasifikasikan jenis penyakit ini berdasarkan gejala-gejala yang ada, sehingga pasien dapat diketahui termasuk dalam skizofrenia jenis apa.

Support Vector Machine (SVM) adalah algoritme *supervised* yang berupa klasifikasi dengan cara membagi data menjadi dua kelas menggunakan garis vektor yang disebut *hyperplane* (Octaviani, et al., 2014). Pada permasalahan yang kompleks atau permasalahan dengan parameter yang banyak, metode ini sangat baik untuk digunakan. Metode ini juga baik digunakan untuk mendiagnosis berbagai macam jenis penyakit. Salah satu kelebihan yang dimiliki metode SVM adalah penanganan *error* pada set data *training* yang menggunakan *Structural Risk Minimization* (SRM). SRM dikatakan lebih baik karena tidak hanya meminimalkan *error* yang terjadi, tetapi meminimalkan faktor-faktor lainnya.

Pada dasarnya metode SVM merupakan metode yang menggunakan *hyperplane* untuk digunakan sebagai pemisah antara data secara linier, sehingga untuk mengatasi permasalahan data yang berupa data nonlinier, maka dapat digunakan teknik *kernel trick*. Metode *Support Vector Machine* (SVM) juga cukup baik dalam menyelesaikan permasalahan klasifikasi *multiclass*. Konsep *One Against All* adalah salah satu cara dalam mengatasi permasalahan *multiclass*. Maka dari itu penelitian ini akan menggunakan algoritme *Support Vector Machine* untuk mengklasifikasikan jenis penyakit skizofrenia.

2. SKIZOFRENIA

Skizofrenia merupakan suatu gangguan terhadap fungsi otak yang merusak cara berpikir seseorang, sehingga tidak terjadi keselarasan antara pola berpikir dengan tingka laku orang tersebut (Hendarsyah, 2016).

Gejala yang diderita pasien Skizofrenia

dikategorikan berdasarkan dampak yang ditimbulkan. Berdasarkan buku pedoman diagnosis yang dipakai oleh para dokter spesialis kejiwaan yaitu PPDGJ jilid 3 (Maslim, 2001), dalam mendiagnosis penyakit skizofrenia pasien dapat dikatakan mengalami penyakit skizofrenia jika mengalami minimal satu gejala yang sangat jelas.

3. METODE PENELITIAN

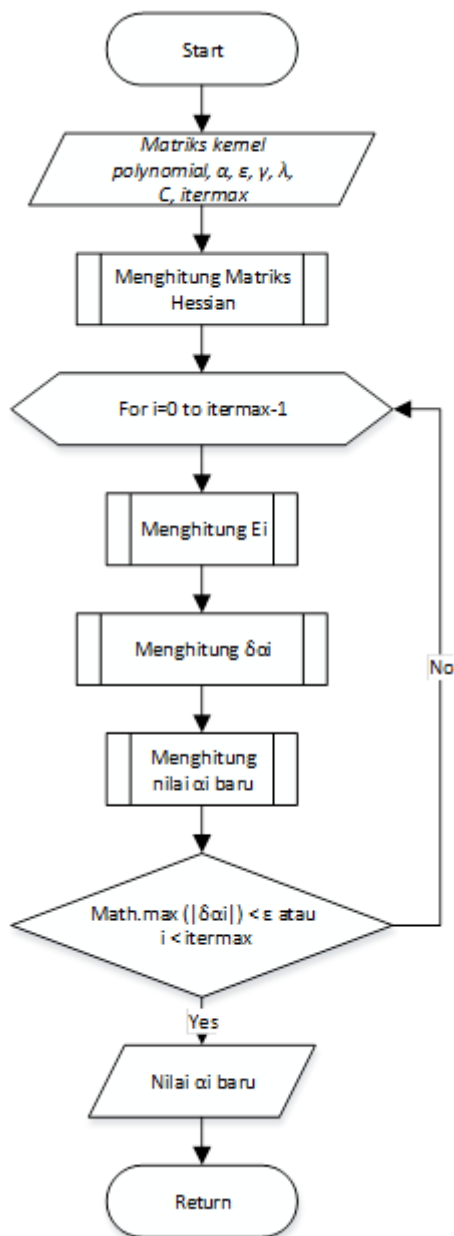
Sistem yang akan dibuat adalah mengenai klasifikasi jenis penyakit Skizofrenia yang terbagi menjadi lima kelas. Kelas 1 adalah jenis Paranoid (F20.0), kelas 2 adalah jenis Hebrefrenik (F20.1), kelas 3 adalah jenis Katatonik (F20.2), kelas 4 adalah jenis Undifferentiated (F20.3), dan kelas 5 adalah jenis Simpleks (F20.6). Data yang digunakan dalam proses klasifikasi adalah data rekam medis dari pasien penyakit Skizofrenia. Data ini akan digunakan sebagai data latih dan data uji. Parameter yang digunakan dalam perhitungan berasal dari gejala-gejala yang ada dalam diagnosis penyakit Skizofrenia.

3.1 Proses Support Vector Machine

Pada langkah awal dalam proses SVM akan dilakukan *input* data dari *database*. Setelah itu data akan dihitung menggunakan *kernel polynomial* untuk diteruskan ke proses *training* dengan menggunakan proses *Sequential Training SVM*. *Output* yang akan dihasilkan adalah berupa hasil klasifikasi, yang berupa perbandingan dari hasil klasifikasi sistem dengan hasil dari pakar.

3.2 Sequential Training SVM

Proses *Sequential Training SVM* adalah proses yang digunakan untuk mencari nilai *hyperplane* terbaik. Proses ini dimulai dari inisialisasi nilai konstanta, perhitungan matriks hessian, perhitungan E_i , $\delta\alpha_i$ hingga menghitung nilai α_i yang baru. Hasil akhir dari proses *training* ini adalah berupa nilai α_i yang akan digunakan untuk proses pengujian. Gambar 1 akan memperlihatkan diagram alir proses SVM.



Gambar 1. Diagram Alir Sequential Training SVM

3.3 Proses Pengujian SVM

Proses pengujian ini akan menggunakan jenis pengujian *K-Fold Cross Validation*, dan akan menggunakan 11 *fold* dalam pengujian ini. Pada proses ini juga akan diterapkan konsep *One Against All* untuk mengatasi permasalahan *multiclass*. Gambar 2 akan menampilkan diagram alir proses pengujian pada SVM.

3.4 One Against All

Penelitian mengenai klasifikasi jenis penyakit skizofrenia membutuhkan konsep *One Against All*, karena jenis penyakit skizofrenia terbagi menjadi 5 kelas. Lalu untuk membedakan setiap kelas maka metode ini

membutuhkan 4 *level* yang berbeda. Pada *level 1* berfungsi untuk memisahkan kelas 1 (*Paranoid*) yang bernilai 1, dengan kelas lainnya yang bernilai -1. Setelah itu pada *level 2* berfungsi untuk memisahkan kelas 2 (*Hebrefenik*) yang bernilai 1, dengan kelas lainnya yang bernilai -1. Lalu pada *level 3* berfungsi untuk memisahkan kelas 3 (*Katatonik*) yang bernilai 1, dengan kelas lainnya yang bernilai -1. Pada *level 4* berfungsi untuk memisahkan kelas (*Undifferentiated*) yang bernilai 1, dengan kelas 5 (*Simpleks*) yang bernilai -1. Selanjutnya untuk menentukan data uji termasuk kelas yang mana, maka data uji akan diproses. Jika nilai data uji bernilai 1 pada *level 1* maka data uji tersebut merupakan kelas 1, dan bila bernilai -1 maka dia akan melanjutkan proses perhitungan kembali ke *level* selanjutnya. Proses ini akan terus diulang hingga data uji bernilai 1 pada *level* tertentu, atau diproses hingga *level* terakhir yaitu *level 4*. Untuk nilai -1 pada *level 4* maka otomatis masuk ke kelas 5. Gambar 3 akan menampilkan diagram alir proses *One Against All*.



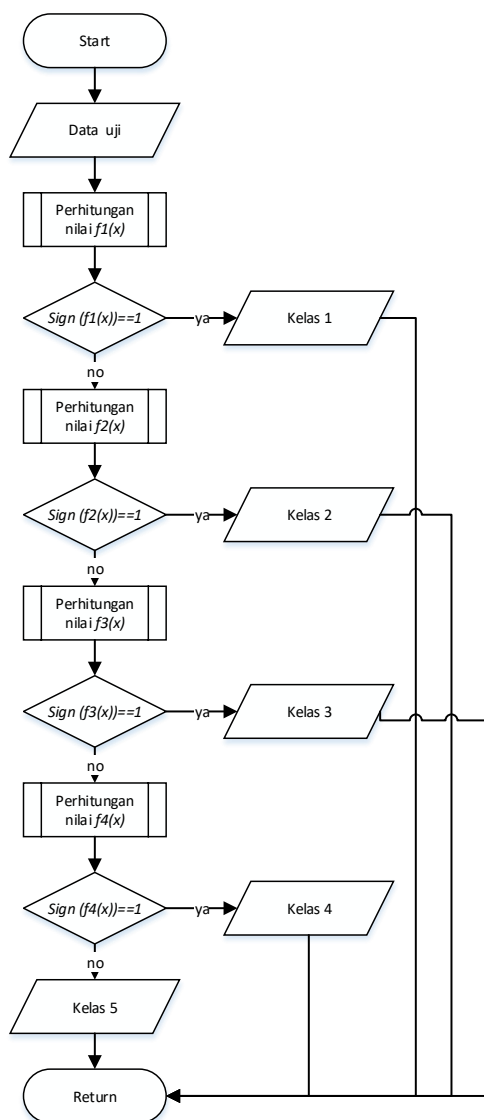
Gambar 2. Diagram Alir Proses Pengujian SVM

4. PENGUJIAN DAN ANALISIS

Pengujian yang akan dilakukan adalah pengujian terhadap nilai dari beberapa macam variabel yang digunakan pada proses perhitungan metode SVM. Proses pengujian ini memakai jenis pengujian *K-Fold Cross Validation* dengan menggunakan *fold* sebanyak 11.

4.1 Pengujian Jenis Kernel

Pengujian ini bertujuan untuk membandingkan nilai akurasi dari setiap jenis *kernel*. Jenis *kernel* yang akan dibandingkan ada tiga jenis yaitu *kernel* linier, *kernel polynomial*, *kernel* RBF. Pada pengujian akan menggunakan beberapa nilai variabel yaitu meliputi nilai $\lambda = 0,01$; $C = 1$;



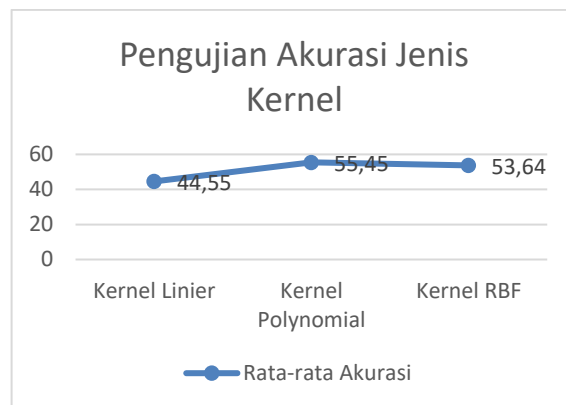
Gambar 3. Diagram Alir One Against All

$\gamma = 0,1$; $itermax = 100$; $\varepsilon = 0,01$. Pada Gambar 4 menunjukkan bahwa *kernel polynomial* memiliki rata-rata tingkat akurasi yang paling tinggi dari jenis *kernel* lainnya yaitu sebesar 55,45%. Hal ini membuktikan bahwa *kernel polynomial* merupakan *kernel* yang cocok untuk menangani permasalahan klasifikasi seperti ini. *Kernel polynomial* merupakan jenis *kernel non-stationary*, dan jenis ini sangat cocok untuk klasifikasi yang memiliki nilai pola yang tidak

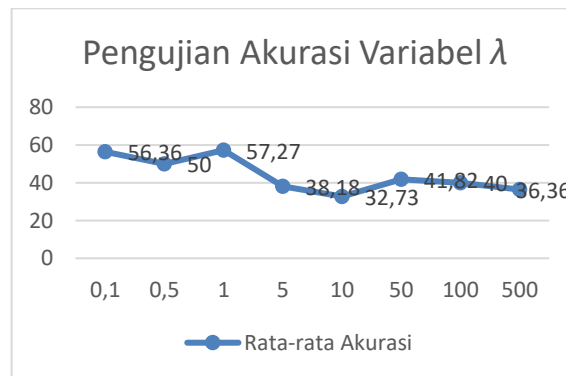
tetap.

4.2 Pengujian Variabel λ (Lambda)

Pengujian ini bertujuan untuk mendapatkan nilai λ yang terbaik dari berbagai jenis nilai *lambda*, mulai dari yang bernilai 0,1 hingga 500. Lalu pengujian ini akan memakai beberapa nilai variabel yaitu meliputi nilai $C = 1$; $\gamma = 0,1$; $itermax = 100$; $\varepsilon = 0,01$. Pengujian ini juga akan menggunakan jenis *kernel polynomial*. Pada Gambar 5 menunjukkan bahwa nilai λ sebesar 1 memiliki tingkat akurasi yang paling tinggi dari nilai lainnya yaitu sebesar 57,27%. Namun secara keseluruhan dapat dilihat bahwa pada nilai λ yang termasuk kecil rata-rata akurasi akan menjadi tinggi, sedangkan pada nilai λ yang tergolong besar tingkat akurasi akan semakin mengecil. Berdasarkan grafik hasil pengujian maka dapat disimpulkan bahwa nilai dari variabel λ yang kecil lebih berpengaruh terhadap tingkat akurasi yang tinggi. Pada proses perhitungan SVM nilai dari variabel λ berfungsi untuk mendapatkan garis pemisah (*hyperplane*) dengan mengatur jarak *margin*.



Gambar 4. Hasil Pengujian Jenis Kernel

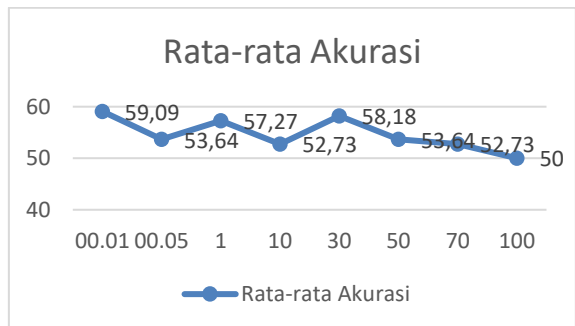


Gambar 5. Hasil Pengujian Variabel λ

4.3 Pengujian Variabel C (Complexity)

Pengujian ini memiliki tujuan untuk

menguji pengaruh variabel C terhadap tingkat akurasi. Terdapat 8 jenis nilai C yang digunakan pada pengujian ini mulai dari yang bernilai 0,1 sampai 100. Ada juga beberapa nilai variabel pendukung yang akan digunakan pada pengujian ini yaitu meliputi nilai $\lambda = 1$; $\gamma = 0,1$; $itermax = 100$; $\varepsilon = 0,01$. *Kernel* yang digunakan pada pengujian variabel ini adalah *kernel polynomial*.



Gambar 6. Hasil Pengujian Variabel C

Pada gambar 6 menunjukkan tingkat akurasi tertinggi didapatkan pada nilai C yang bernilai 0,1 yaitu dengan tingkat akurasi sebesar 59,09%. Pengujian ini memiliki kesamaan dengan hasil pengujian nilai λ yaitu terjadi ketidakstabilan pada hasil akurasi dari beberapa nilai C yang diuji. Pada proses perhitungan SVM nilai dari variabel C berfungsi untuk mengurangi nilai *error* pada *set data training*. Nilai dari variabel C berbanding lurus dengan nilai *error*, yaitu jika ingin mendapatkan nilai *error* yang kecil maka nilai variabel C juga harus kecil. Nilai C juga berpengaruh terhadap waktu komputasi pada proses perhitungan data *training*. Semakin besar nilai C maka waktu komputasi semakin lama.

4.4 Pengujian Variabel γ (Gamma)

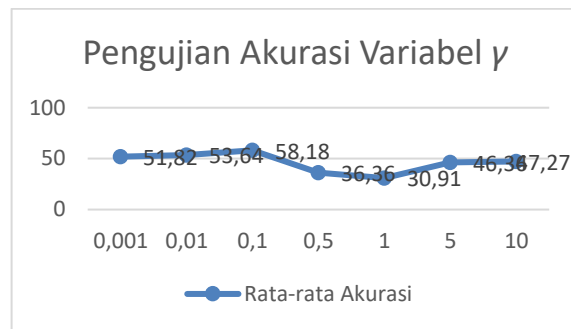
Pengujian ini bertujuan untuk mengetahui seberapa besar pengaruh nilai γ terhadap tingkat akurasi sistem. Pada pengujian ini terdapat beberapa nilai γ yang akan diuji yaitu ada 7 jenis nilai mulai dari yang bernilai 0,001 sampai 10. Pada pengujian akan menggunakan beberapa nilai variabel yaitu meliputi nilai $\lambda = 1$; $C = 0,1$; $itermax = 100$; $\varepsilon = 0,01$. Pengujian ini juga akan menggunakan jenis *kernel polynomial*.

Pada Gambar 7 menunjukkan bahwa tingkat akurasi tertinggi terjadi pada nilai γ yang bernilai 0,1 dengan nilai akurasi sebesar 58,18%. Dapat dilihat pada pengujian ini, bahwa tingkat akurasi cenderung menurun pada nilai γ yang semakin besar. Pada dasarnya fungsi nilai γ adalah untuk mengatur laju pembelajaran (*learning rate*). Hal inilah yang menyebabkan jika nilai γ semakin

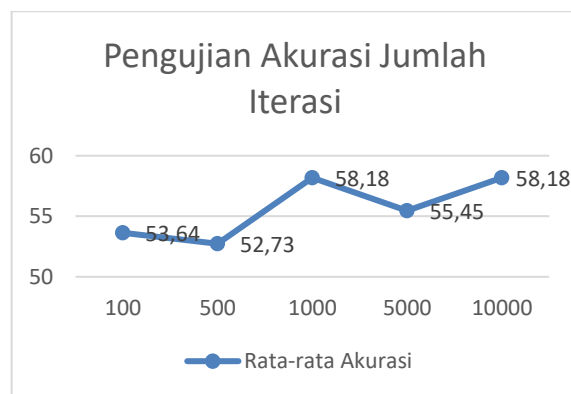
tinggi maka nilai *learning rate* akan semakin tinggi, dan jika nilai *learning rate* semakin tinggi maka tingkat ketelitian pada sistem ini akan berkurang.

4.5 Pengujian Jumlah Iterasi

Pengujian ini memiliki tujuan untuk menguji pengaruh jumlah iterasi terhadap tingkat akurasi.



Gambar 7. Hasil Pengujian Variabel γ



Gambar 8. Hasil Pengujian Jumlah Iterasi

Terdapat beberapa jumlah iterasi yang digunakan pada pengujian ini mulai dari yang bernilai 100 sampai 10000. Ada juga beberapa nilai variabel pendukung yang akan digunakan pada pengujian ini yaitu meliputi nilai $\lambda = 1$; $\gamma = 0,1$; $C = 0,1$; $\varepsilon = 0,01$. *Kernel* yang digunakan pada pengujian variabel ini adalah *kernel polynomial*.

Pada Gambar 8 menunjukkan tingkat akurasi tertinggi didapatkan pada jumlah iterasi yang bernilai 1000 dan 10000 yaitu dengan tingkat akurasi sebesar 58,18%. Pengaruh jumlah iterasi pada proses pelatihan ini yaitu perubahan nilai α yang merupakan nilai *support vector*. Pada dasarnya tingkat akurasi akan semakin kecil jika jumlah iterasi terlalu banyak, hal ini disebabkan karena nilai *support vector* menjadi tidak seimbang dan juga data akan menjadi berjauhan terhadap garis *hyperplane* (Hasanah, Resita M.,

Pratama, & Cholissodin, 2016)

4.6 Analisis Data

Pada bagian ini akan dijelaskan mengenai pengaruh data terhadap hasil akurasi. Terdapat beberapa faktor yang menyebabkan nilai akurasi menjadi tidak maksimal. Faktor yang pertama adalah jumlah data yang tidak rata antar tiap kelas, contohnya adalah data pada kelas 1 berjumlah 32, data pada kelas 2 berjumlah 30, data pada kelas 3 berjumlah 14 data, data pada kelas 4 berjumlah 19, sedangkan data pada kelas 5 berjumlah 16 data. Hal ini menyebabkan data yang digunakan sebagai data pembelajaran tidak seimbang karena salah satu kelas memiliki data lebih banyak. Faktor yang kedua adalah variasi pola data pada satu kelas sangat beragam (non-stationary) dan tidak memiliki pola yang tetap.

5. KESIMPULAN

Agoritme *Support Vector Machine* (SVM) dapat diterapkan pada klasifikasi jenis penyakit skizofrenia, dan jenis penyakit ini dibagi kedalam lima jenis penyakit. Hasil pengujian yang didapatkan dengan melakukan pengujian *K-Fold Cross Validation*, dan fold yang digunakan sebanyak 11 fold. Tingkat akurasi tertinggi yang didapatkan pada penelitian ini adalah sebesar 59,09% dengan nilai variabel $\lambda = 1$; $C = 0,1$; $\gamma = 0,1$; $itermax = 100$; $\varepsilon = 0,01$; dan juga menggunakan *kernel polynomial*. Hasil akurasi pada penelitian terbilang cukup rendah. Hal yang menyebabkan akurasi rendah pada penelitian ini yaitu karena data yang digunakan tidak rata pada setiap kelas, dan juga pola pada setiap kelas berbeda-beda sehingga sulit untuk menentukan pola yang terbaik.

DAFTAR PUSTAKA

- Hasanah, U., Resita M., L., Pratama, A. & Cholissodin, I., 2016. Perbandingan Metode SVM, Fuzzy K-NN, Dan BDT-SVM Untuk Klasifikasi Detak Jantung Hasil Elektrokardiografi. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 3(3), pp. 201-207.
- Hendarsyah, F., 2016. Diagnosis dan Tatalaksana Skizofrenia Paranoid dengan Gejala-Gejala Positif dan Negatif. *J Medula Unila*, 4(3), p. 58.
- Maryatun, S., 2015. Peningkatan Kemandirian Perawatan Diri Pasien Skizofrenia Melalui Rehabilitasi Terapi Gerak. *Jurnal*

Keperawatan Sriwijaya, 2(2).

- Maslim, R., 2001. Buku Saku Diagnosis Gangguan Jiwa : Rujukan Ringkas Dari PPDGJ-III. 1 penyunt. Jakarta: Bagian Ilmu Kedokteran Jiwa FK-UNIKA Atma Jaya.
- Octaviani, P. A., Wilandari, Y. & Ispriyanti, D., 2014. Penerapan Metode Klasifikasi Support Vector Machine (SVM) Pada Data Akreditasi Sekolah Dasar (SD) Di Kabupaten Magelang. *Jurnal Gaussian*, 3(4), pp. 811-820.