

Implementasi *Support Vector Machine* (SVM) dalam Memprediksi Jumlah Penyakit Demam Berdarah

¹Danu Sasmita, ²Favorisen Rosyking Lumbanraja, ³Kurnia Muludi, dan ⁴Astria Hijriani

^{1,2,3,4} Jurusan Ilmu Komputer Fakultas Matematika Ilmu Pengetahuan Alam, Universitas Lampung
Jalan Soemantri Brojonegoro No.1 Gedung Meneng, Bandar Lampung, Provinsi Lampung, Indonesia
e-mail : ¹ sasmitadanu@gmail.com, ² favorisen.lumbanraja@fmipa.unila.ac.id, ³ kurnia@unila.ac.id,
⁴ astria.hijriani@fmipa.unila.ac.id

Abstract — *Dengue Hemorrhagic Fever (DHF) is an infectious disease caused by dengue virus infection and is transmitted through the bite of female mosquito species Aedes aegypti and Aedes albopictus. Environmental factors are one of the causes of the high prevalence of dengue fever, including the layout of buildings, water reservoirs, indentations in the soil, temperature and other things that can help the Aedes mosquito life cycle take place. The purpose of this study is to predict the spread of dengue disease using the SVM (Support Vector Machine) method with rainfall data in Singapore from 2014 to 2018, weather data and pain data, comparing this study with previous research by Adeline Ong in 2014 entitled "Predicting Dengue Cases in Singapore", as well as knowing the results of predicting the distribution of DHF using the SVM method in the form of variance values (R^2) with linear, gaussian and polynomial kernels. The results of the experiment found that the lowest error value was shown by the linear kernel with an error rate of 35.15%, with a variance value of 64.85%.*

Keywords: *Dengue Fever; K-Fold Cross Validation; Support Vector Machine.*

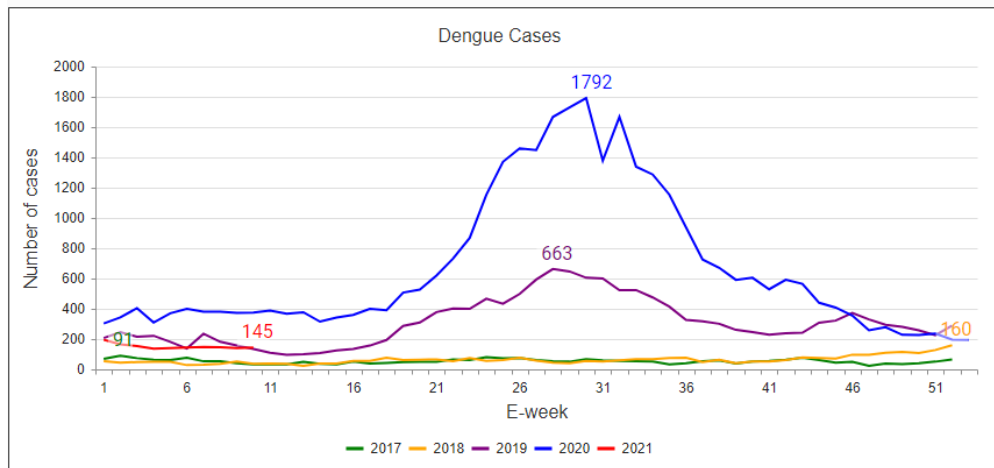
1. PENDAHULUAN

Penyakit Demam Berdarah Dengue (DBD) merupakan salah satu penyakit menular yang disebabkan oleh salah satu infeksi virus dengue melalui gigitan nyamuk betina spesies *Aedes aegypti* dan *Aedes albopictus* [1]. Proses penularan virus dengue dari nyamuk aedes betina terjadi pada saat mulut nyamuk menusuk bagian kulit manusia sehingga menyebabkan virus bertransmisi [2].

Nyamuk betina berjenis aedes dapat dengan mudah menjumpai manusia karena hanya beraktifitas di wilayah yang memiliki ketinggian kurang dari 1 kilometer diatas permukaan laut, artinya dapat hidup di daerah perkotaan atau pedesaan dengan dataran rendah [1]. DBD banyak ditemukan pada daerah yang beriklim subtropis dan tropis, dengan variasi lokal dalam risiko dipengaruhi oleh curah hujan, suhu dan urbanisasi yang tidak terencana [3].

Menurut World Health Organization (WHO) pada tahun 2017, penyakit DBD pertama kali ditemukan di Filipina dan Thailand pada tahun 1950an dan telah mewabah di seluruh dunia dalam beberapa dekade terakhir [4]. Salah satu perkiraan terakhir menunjukkan bahwa 390 juta jiwa di 128 negara, beresiko untuk terinfeksi virus dengue. Pada tahun 2010, tercatat hampir 2,2 juta kasus yang dilaporkan dan meningkat menjadi 3,2 juta pada tahun 2015.

Tahun 1953 merupakan tahun dimana kasus demam berdarah pertama kali ditemukan di Filipina, sejak saat itu beberapa negara anggota WHO khususnya yang berada di Pasifik Barat dan Asia Tenggara secara rutin membuat laporan mengenai kasus demam berdarah yang terjadi pada negara masing-masing, salah satunya Singapura. Berdasarkan informasi dari National Environment Agency (NEA) atau dalam Bahasa Indonesia Badan Lingkungan Nasional Singapura, kejadian demam berdarah di Negara Singapura selalu mengalami kenaikan yang cukup tinggi dari tahun 2017 hingga tahun 2020 [5]. Berikut Gambar 1 menunjukkan kasus demam berdarah di Singapura dari Minggu pertama tahun 2017 hingga minggu ke 10 tahun 2020 [5].



Gambar 1. Kasus demam berdarah di Negara Singapura [5].

Gambar 1 menunjukkan kasus demam berdarah di Negara Singapura selama 51 minggu setiap tahunnya. Dimulai dari tahun 2017, memiliki 91 kasus demam berdarah yang merupakan kasus terkecil yang ditunjukkan diagram tersebut. Tahun 2018 kasus demam berdarah mencapai 160 kasus, artinya kasus demam berdarah naik sekitar 75% dari tahun sebelumnya. Tahun 2019 hal mengejutkan terjadi, karena kasus demam berdarah naik sebesar 310% dari tahun sebelumnya, yakni terdapat 663 kasus demam berdarah pada tahun 2019. Tahun 2020 kasus demam berdarah juga melambung tinggi mencapai angka 1792 kasus [5].

Faktor lingkungan merupakan salah satu penyebab tingginya kasus penyebaran penyakit demam berdarah, diantaranya tata letak bangunan, penampungan air, lekukan pada tanah, suhu serta hal-hal lain yang dapat membantu berlangsungnya siklus hidup nyamuk aedes. Jenjang Pendidikan juga dapat menjadi salah satu faktor penyebab terjadinya demam berdarah, karena merupakan penentu bagi suatu individu untuk memilah tepat atau tidaknya informasi yang diterima mengenai pengetahuan umum demam berdarah, mengenali gejala demam berdarah, menerapkan hal yang dapat meminimalisir faktor terjadinya kasus demam berdarah [6].

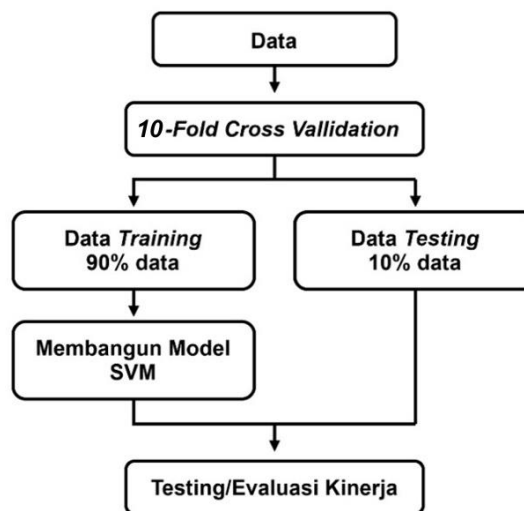
Untuk mengetahui seberapa besar potensi persebaran penyakit demam berdarah di suatu wilayah, perlu dilakukan adanya peramalan berupa klasifikasi penyebaran penyakit demam berdarah di wilayah tersebut. Hasil dari klasifikasi diharapkan dapat mempermudah untuk menentukan langkah yang tepat dalam proses penanggulangan penyakit demam berdarah, sehingga dapat meningkatkan derajat kesehatan masyarakat. Salah satu metode peramalan yang baik ialah Support Vector Machine (SVM), hal ini dikarenakan SVM merupakan metode statistik yang memiliki kemampuan penalaran yang mampu mengklasifikasikan apakah suatu pola masuk kedalam data yang digunakan pada fase pembelajaran atau tidak [7].

Tujuan dari penelitian ini yaitu melakukan prediksi persebaran penyakit DBD menggunakan metode SVM dengan data curah hujan di Singapura tahun 2014 hingga 2018, data cuaca dan data kesakitan, membandingkan penelitian ini dengan penelitian sebelumnya oleh Adeline Ong pada tahun 2014 yang berjudul "Predicting Dengue Cases in Singapore", serta mengetahui hasil prediksi persebaran penyakit DBD menggunakan metode SVM berupa nilai varian (R^2) dengan kernel linear, gaussian dan polynomial.

2. METODE PENELITIAN

2.1. Alur Metode Penelitian

Metode yang dilakukan dalam melakukan penelitian tentang prediksi persebaran penyakit demam berdarah di Singapura menggunakan metode SVM dilakukan dengan beberapa tahap, tahapan tersebut dapat dilihat pada Gambar 2.



Gambar 2. Tahapan metode SVM dengan *k-fold cross validation*.

Penjelasan dari tahapan metode penelitian pada Gambar 2 adalah sebagai berikut:

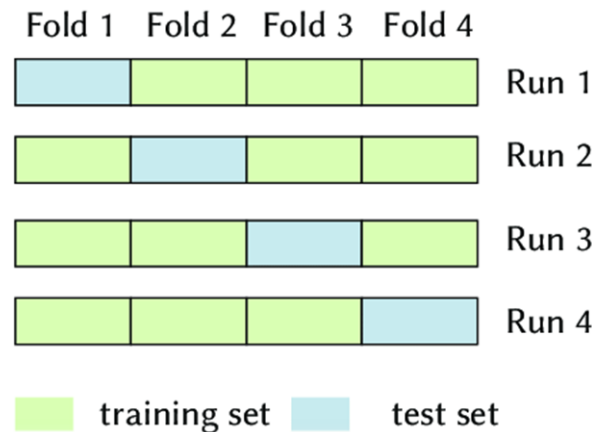
2.1.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data yang digunakan oleh Adeline Ong pada penelitiannya yang berjudul “Predicting Dengue Cases in Singapore” pada tahun 2019 yang diperoleh dari situs <https://github.com/adelweiss/Dengue/tree/master>. Data yang digunakan meliputi 15 variabel diantaranya dengue_cases, rainfall, mean_temp, max_temp, fever, nausea, headache, ache_pain, eye_pain, dengue, rashes, vomiting, dengue_cases, temp_range.

2.1.2. K-Fold Cross Validation

Pada proses ini data dibagi menjadi dua bagian untuk proses testing dan training. Training data yang digunakan sebanyak 90%, sedangkan pada proses testing data yang digunakan hanya sebanyak 10%. Setiap titik data setidaknya memiliki peluang untuk divalidasi, maka set training dan validation harus cross-over dalam putaran berturut-turut [8].

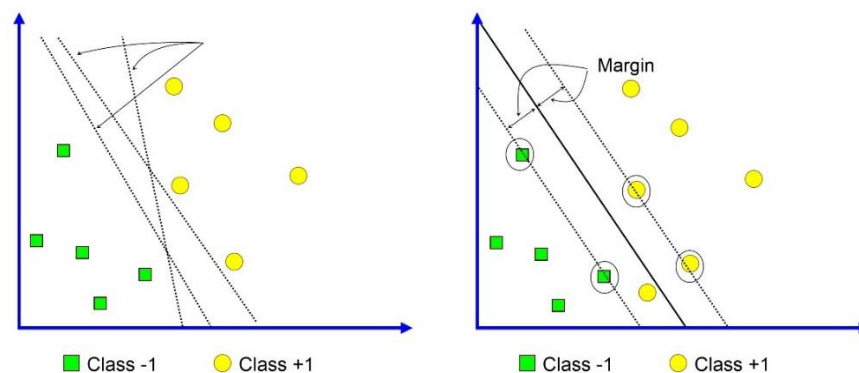
K-fold cross-validation merupakan bentuk dasar dari *cross-validation*. Dalam *k-fold cross-validation*, dilakukan partisi data menjadi *k* atau segmen yang berukuran sama atau hampir sama. Selanjutnya dilakukan pelatihan iterasi dan validasi hingga dalam setiap iterasi lipatan data yang berbeda diulurkan untuk validasi, sedangkan *k-1* lipatan yang tersisa digunakan untuk pembelajaran [9]. Ilustrasi mengenai cara kerja *K-fold cross-validation* dapat dilihat pada Gambar 3.



Gambar 3. Ilustrasi cara kerja *k-fold cross validation* [10].

2.1.3. Pemodelan dengan SVM

Untuk melakukan prediksi, terlebih dahulu menentukan model svm terbaik dari data training yang telah di training dengan model SVM, dengan terlebih dahulu menentukan tipe kernel dan nilai parameter.



Gambar 4. Ilustrasi SVM dalam mencari *hyperplane*.

Pada ilustrasi Gambar 4 terdapat dua bagian class yang telah dipisahkan oleh *hyperplane*, yaitu kelas negative yang diilustrasikan pada kelompok berwarna merah yang bernilai -1 [11]. Berikut persamaan (1) :

$$X_i \cdot W + b \leq -1 \text{ untuk } Y_i = -1 \quad (1)$$

Serta kelas positif yang diilustrasikan pada kelompok berwarna kuning yang bernilai +1 [11]. Juga dapat dilihat pada persamaan (2) :

$$X_i \cdot W + b \geq 1 \text{ untuk } Y_i = 1 \quad (2)$$

Variabel X_i adalah variabel yang ditunjukkan untuk data ke- i pada dataset, dan Y_i merupakan variabel yang ditunjukkan untuk kelas data ke- i pada dataset. Sedangkan W menunjukkan nilai pada *support vector* yang keberadaannya tegak lurus terhadap *hyperplane* dan biasa disebut dengan bobot *vector*.

2.1.4. Evaluasi

Model SVM yang telah didapat dari proses pemodelan, dilakukan pengujian dengan data testing yang dari proses pembagian data set. Apabila data yang digunakan terbatas, maka digunakan teknik *k-fold cross validation*.

3. TAHAPAN PENGUJIAN

Dalam penelitian ini percobaan prediksi dilakukan menggunakan 3 *kernel*, yaitu *kernel linear*, *gaussian*, dan *polynomial*. Sebelum melakukan prediksi menggunakan model SVM dengan ketiga *kernel* tersebut, terlebih dahulu ada beberapa hal yang harus dilakukan diantaranya import data, pemrosesan data, pembagian data set, penentuan *k-fold cross validation*, dan membuat model regresi.

3.1. Import Data

Data dengan format *Comma Separated Values* (CSV) diolah menggunakan R Studio, data yang telah dimasukkan dapat terbaca didalam R Studio dan didapat sebanyak 248 data dengan 15 variabel yang terdiri dari: *dengue_cases*, *rainfall*, *mean_temp*, *max_temp*, *min_temp*, *fever*, *nausea*, *headache*, *ache_pain*, *eye_pain*, *dengue*, *rashes*, *vomiting*, *dengue_cases_+1*, *temp_range*.

3.2. Pemrosesan Data

Untuk memudahkan dalam proses percobaan, maka data terlebih dahulu diproses dengan mengubah nama set data agar lebih mudah digunakan. Dalam proses ini nama set data diubah menjadi "dengue".

3.3. Pembagian Data Set

Proses ini dilakukan dengan membagi data yang sudah ada menjadi 2 bagian, yaitu bagian untuk data latih dan bagian untuk data uji. Proses pembagian data dilakukan dengan metode *10-Fold Cross Validation*, sebanyak 10% data dimasukkan kedalam data uji dan sebanyak 90% data dimasukkan kedalam data latih. Kemudian didapat data sebanyak 25 untuk data uji dan data sebanyak 223 untuk data latih.

3.4. Membuat Model Regresi

Model regresi dilakukan menggunakan metode SVM *linear* dengan *kernel linear*, *gaussian*, dan *polynomial* serta *K-Fold Cross Validation* sebanyak *10-Fold*.

3.5. Prediksi dengan Model SVM Menggunakan *Kernel Linear*, *Gaussian* dan *Polynomial*.

Prediksi adalah proses inti dari penelitian ini, yang pertama proses prediksi dilakukan dengan data latih yang tadi telah dibagi sebanyak 90% dari data yang ada. Kemudian langkah selanjutnya yaitu menemukan nilai varian dari proses prediksi berupa *Sum of square error* (SSE), *total sum of square* (SST), dan R^2 . Percobaan untuk mencari nilai varian dilakukan sebanyak 10 kali, karena menggunakan *10-Fold Cross Validation*.

Evaluasi pengujian menggunakan SVM dengan *kernel linear* dapat dilihat pada Tabel 1.

Tabel 1. Hasil menggunakan SVM dengan *kernel linear*

Fold	Evaluasi Pengujian <i>Kernel Linear</i>		
	SSE	SST	R^2
1	3220.082	12341.04	0.7754638
2	2707.083	4961.04	0.4543315
3	3876.384	4535.44	0.1453124
4	4205.452	6176	0.3190655

Fold	Evaluasi Pengujian <i>Kernel Linear</i>		
	SSE	SST	R ²
5	11296.94	144442	0.9217891
6	51717.66	662578	0.9219448
7	23911.97	77925.83	0.6931445
8	37462.68	64651.44	0.4205437
9	36729.51	822486.2	0.9553433
10	84169.94	692432	0.878443
Rata-rata	25929.7701	249252.899	0.64853816

Dari 10 *fold* menggunakan metode SVM dengan *kernel linear* didapat rata-rata nilai SSE pada pengujian ini bernilai 25929.77, dan rata-rata nilai SST bernilai 249252.89. Sedangkan nilai varian (R²) sebesar 0.6485 jika dikonversi dalam bentuk persen berarti sebesar 64.85%.

Setelah nilai varian ditemukan, kita dapat mengetahui nilai tingkat kesalahan dari percobaan menggunakan metode SVM dengan *kernel linear*, yaitu 100% dikurangi 64.85% sehingga didapat nilai tingkat kesalahan dari percobaan menggunakan metode SVM dengan *kernel linear* yaitu sebesar 35.15%.

Evaluasi pengujian menggunakan SVM dengan *kernel gaussian* dapat dilihat pada Tabel 2.

Tabel 2. Hasil menggunakan SVM dengan *kernel gaussian*

Fold	Evaluasi Pengujian <i>Kernel Gaussian</i>		
	SSE	SST	R ²
1	5509.265	14341.04	0.6158393
2	4954.671	4961.04	0.001283785
3	3877.98	4535.44	0.1449606
4	5662.593	6176	0.08312938
5	19371.84	144442	0.865885
6	96920.23	662578	0.8537225
7	38801.22	77925.83	0.502075
8	43457.31	64651.44	0.3278214
9	99970.04	922486.2	0.8784538
10	242806.8	692432	0.6493421
Rata-rata	56133.1949	259452.899	0.492251287

Pengujian pada *kernel gaussian* dengan 10 *fold* menggunakan metode SVM didapat rata-rata nilai SSE pada pengujian ini bernilai 56133.19, dan rata-rata nilai SST bernilai 259452.89. Sedangkan nilai varian (R²) sebesar 0.4922 jika dikonversi dalam bentuk persen berarti sebesar 49.22%.

Dapat diketahui nilai tingkat kesalahan dari percobaan menggunakan metode SVM dengan *kernel gaussian* berdasarkan nilai varian yang telah didapat, yaitu 100% dikurangi 49.22% sehingga didapat nilai tingkat kesalahan dari percobaan menggunakan metode SVM dengan *kernel gaussian* yaitu sebesar 50.78%. Jika dibandingkan dengan *kernel linear*, tingkat kesalahan pada *kernel gaussian* memiliki nilai tingkat kesalahan yang cukup tinggi.

Evaluasi pengujian menggunakan SVM dengan *kernel Polynomial* dapat dilihat pada Tabel 3.

Tabel 3. Hasil menggunakan SVM dengan *kernel polynomial*

Fold	Evaluasi Pengujian <i>Kernel Polynomial</i>		
	SSE	SST	R ²
1	3542.553	14341.04	0.7529779
2	2639.198	4961.04	0.4680152
3	3820.649	4535.44	0.1576013
4	5024.96	6176	0.1863731
5	12624.68	144442	0.9125969
6	53491.07	662578	0.9192683
7	23732.92	77925.83	0.6954422
8	37902.33	64651.44	0.4137435
9	34831.22	822486.2	0.9576513
10	84816.93	692432	0.8775087
Rata-rata	26242.651	249452.899	0.63411784

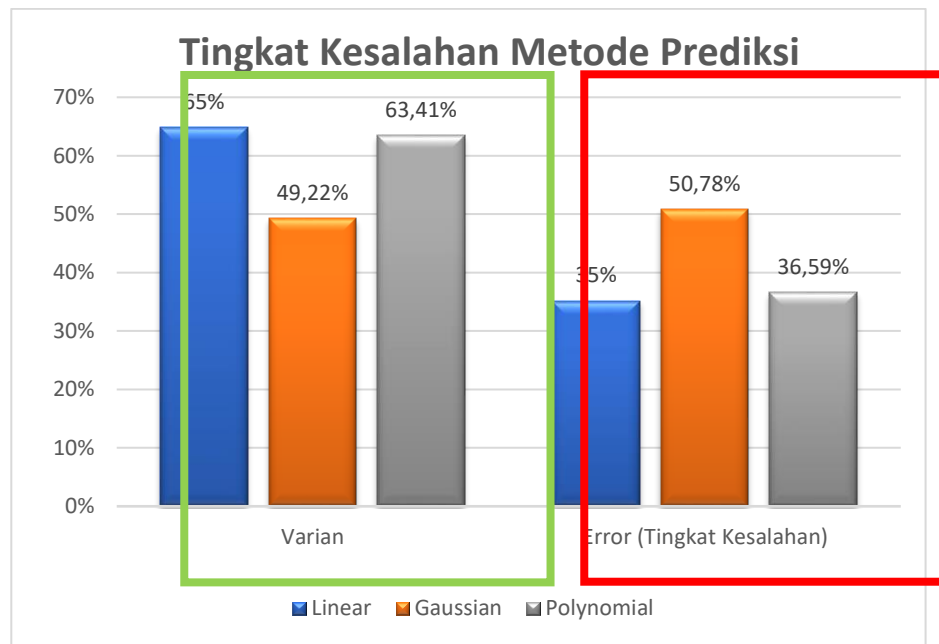
Dengan *kernel polynomial* pada 10 *fold* menggunakan metode SVM didapat rata-rata nilai SSE pada pengujian ini bernilai 26242.65, dan rata-rata nilai SST bernilai 249452.89. Sedangkan nilai varian (R²) sebesar 0.6341 jika dikonversi dalam bentuk persen berarti sebesar 63.41%.

Tingkat kesalahan dari percobaan menggunakan metode SVM dengan *kernel polynomial* dapat dihitung ketika nilai varian telah didapat, yakni 100% dikurangi 63.41% sehingga didapat nilai tingkat kesalahan dari percobaan menggunakan metode SVM dengan *kernel polynomial* yaitu sebesar 36.59%.

3.6. Hasil Pengujian

Pada penelitian ini dilakukan percobaan prediksi menggunakan metode SVM dengan *kernel linear*, *gaussian* dan *polynomial*. Dengan data yang diambil di distrik Changi, Singapura, dalam penelitian ini terdapat sebanyak 248 data dengan variable sebanyak 15 variabel yang terdiri dari 2 variable kasus DBD yaitu *dengue_case* dan *dengue_case_+1*, 5 variable cuaca seperti *rainfall*, *mean_temp*, *max_temp*, *temp_range* dan *min_temp*, serta 8 variable kesakitan yang terdiri dari *fever*, *nausea*, *headache*, *ache_pain*, *eye_pain*, *dengue*, *rashes* dan *vomiting*.

Berdasarkan 15 variabel yang tersebut, 1 variabel digunakan sebagai variabel kriterium atau variabel yang dijadikan sebagai objek prediksi yaitu variabel *dengue_case*, dan 14 variabel lainnya dijadikan sebagai variabel predictor yaitu variabel *dengue_case_+1*, *rainfall*, *mean_temp*, *max_temp*, *temp_range*, *min_temp*, *fever*, *nausea*, *headache*, *ache_pain*, *eye_pain*, *dengue*, *rashes* dan *vomiting*. Variabel *dengue_case* yang dijadikan sebagai objek prediksi merupakan data kasus DBD perminggu di distrik Changi, Singapura dari tahun 2014 hingga 2018. Hasil pengujian menggunakan metode SVM dengan *kernel linear*, *gaussian* dan *polynomial* menghasilkan perbandingan tingkat kesalahan model dari masing-masing *kernel* yang dapat dilihat pada Gambar 3.



Gambar 5. Grafik tingkat kesalahan metode prediksi

Grafik yang berada di didalam kotak berwarna hijau menunjukkan grafik nilai varian dari kernel *linear*, *gaussian* dan *polynomial*. Sedangkan grafik yang berada didalam kotak berwarna merah menunjukkan nilai tingkat kesalahan dari metode prediksi yang dilakukan menggunakan kernel *linear*, *gaussian* dan *polynomial*.

Sehingga dari grafik tersebut tentu dapat ditarik kesimpulan bahwa semakin besar nilai varian, maka semakin kecil tingkat kesalahan dari model prediksi, yang artinya semakin baik model prediksi yang digunakan. Dilihat dari tingkat kesalahannya, model prediksi paling baik terdapat pada *kernel Linear* dengan nilai error 35%, sedangkan model prediksi dengan tingkat kesalahan paling tinggi terdapat pada *kernel gaussian* dengan nilai error sebesar 50.78%.

Percobaan dari *kernel linear*, *gaussian* dan *polynomial* memiliki waktu percobaan yang berbeda pada setiap kernel, hal ini dapat dilihat pada Tabel 4.

Tabel 4. Lama waktu percobaan

Kernel	Running Time
<i>Linear</i>	22.22 secs
<i>Gaussian</i>	1.05 mins
<i>Polynomial</i>	10 days

Perbandingan lama waktu percobaan pada tabel tersebut dapat dilihat bahwa waktu paling cepat terdapat pada *kernel linear* yaitu 22.22 detik, sedangkan waktu percobaan paling lama didapat pada *kernel polynomial* yang memerlukan waktu percobaan hingga 10 hari.

Jika dilihat dari dua indikator tersebut yaitu nilai error dan lama waktu percobaan yang diperlukan masing-masing *kernel*, metode terbaik dalam penelitian ini didapat pada *kernel linear* dengan nilai error paling rendah yaitu sebesar 35% dan hanya memerlukan waktu percobaan sekitar 22.22 detik, waktu ini paling cepat diantara *kernel* lainnya. Sedangkan hasil penelitian paling buruk didapat pada *kernel gaussian* dengan nilai error paling tinggi yaitu sebesar 50.78% dengan waktu percobaan sekitar 1.05 menit.

4. KESIMPULAN

Berdasarkan percobaan yang dilakukan didapatkan kesimpulan sebagai berikut:

1. Dari percobaan prediksi menggunakan *kernel linear*, *gaussian* dan *polynomial*, nilai error paling rendah didapat pada *kernel linear* yaitu 35.15%, mengingat nilai varian pada *kernel linear* paling tinggi dari *kernel* lainnya yaitu sebesar 64.85%.
2. Nilai error paling buruk didapat pada *kernel gaussian* yaitu 50.78%, hal ini dikarenakan *kernel gaussian* memiliki nilai varian paling rendah yaitu 49.22%.
3. Dari pernyataan pada nomor 1 dan 2, dapat disimpulkan bahwa semakin besar nilai varian maka semakin rendah nilai error yang artinya menunjukkan bahwa semakin baik metode yang digunakan.
4. Waktu percobaan prediksi pada *kernel linear* membutuhkan waktu paling lama dari kedua *kernel* lainnya yaitu selama 10 hari. Sedangkan *kernel linear* hanya membutuhkan waktu 22.22 detik, dan *kernel polynomial* membutuhkan waktu 1.05 menit.
5. Jika dilihat dari nilai error dan waktu percobaan, percobaan paling baik didapat pada *kernel linear*, dengan nilai error paling rendah sekaligus waktu percobaan paling cepat dari *kernel* lainnya.
6. Dibandingkan dengan penelitian oleh Adeline Ong pada tahun 2019 yang memiliki nilai error hanya sebesar 23%, penelitian ini memiliki nilai error sebesar 35% pada *kernel linear* yang artinya tingkat kesalahan pada penelitian ini lebih tinggi dari penelitian tersebut.

DAFTAR PUSTAKA

- [1] A. Sukohar, "Fakultas Kedokteran Universitas Lampung Demam Berdarah Dengue (DBD) Fakultas Kedokteran Universitas Lampung", *Medula*, Vol 2 No 2, 1–15, 2014.
- [2] I. D. Manuaba, I. W. P. Sutirtayasa, & D. R. Dewi, "Immunopatogenesis infeksi virus dengue", *E-Jurnal Medika Udayana*, Vol 2 No 2, 1–12, 2013.
- [3] A. Candra, "Demam Berdarah Dengue : Epidemiologi, Patogenesis, dan Faktor Risiko Penularan", *Jurnal Aspirator* Vol 2 No 2, 110–119, 2010.
- [4] (2019) Apps.who.int website. [Online]. Available: <https://apps.who.int/mediacentre/factsheets/fs117/en/index.html>
- [5] (2020) Nea.goc.sg website. [Online]. Available: <https://www.nea.gov.sg/dengue-zika/dengue/dengue-cases>
- [6] R. D. Prasetyani, "Faktor-Faktor Yang Berhubungan Dengan Kejadian Demam Berdarah Dengue," *Journal Majority*, Vol 4 No 7, 61–66, 2015.
- [7] A. S. Nugroho, A. B. Witarto, & D. Handoko, "Support Vector Machine Teori dan Aplikasinya dalam Bioinformatika1", *Proceedings of the 2011 Chinese Control and Decision Conference*, 842–847, 2011.
- [8] B. Santoso, *Data Mining, Teori dan Aplikasi*, Jakarta: Graha Ilmu, 2007.
- [9] R. Kohavi, & F. Provost, "Glossary of terms: Special Issue on Applications of Machine Learning and the Knowledge Discovery Process", *Machine Learning*, Vol 30 No 2, 271-274, 1998.
- [10] (2019) Researchgate.net website. [Online]. Available: https://www.researchgate.net/figure/The-technique-of-KFold-cross-validation-illustrated-here-for-the-case-K-4-involves_fig10_278826818
- [11] I. M. Parapat, "Penerapan Metode Support Vector Machine (Svm) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak", Program Studi Teknik Informatika, Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya, Malang, 2017