

**PREDIKSI JUMLAH PENYAKIT DEMAM BERDARAH
MENGUNAKAN METODE *SUPPORT VECTOR MACHINE*
(STUDI KASUS PENYEBARAN DBD DI SINGAPURA)**

(Skripsi)

Oleh

DANU SASMITA

NPM 1617051133



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022**

**PREDIKSI JUMLAH PENYAKIT DEMAM BERDARAH
MENGUNAKAN METODE *SUPPORT VECTOR MACHINE*
(STUDI KASUS PENYEBARAN DBD DI SINGAPURA)**

Oleh

DANU SASMITA

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA ILMU KOMPUTER

Pada

**Jurusan Ilmu Komputer
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Lampung**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022**

ABSTRACT

PREDICTION OF THE NUMBER OF DENGUE FEVER DISEASES USING THE SUPPORT VECTOR MACHINE METHOD (CASE STUDY OF THE SPREAD OF DHF IN SINGAPORE)

By

Danu Sasmita

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 DHF, 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. Forecasting in the form of a classification of the spread of dengue fever is needed to find out how big the potential spread of dengue fever is in an area which is expected to make it easier to determine the right steps in the process of controlling dengue fever, so as to improve the health status of the community. The purpose of this study was to predict the spread of dengue fever using the Support Vector Machine (SVM) 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 Fever Cases in Singapore", and knowing the results of predicting the distribution of DHF using the SVM method in the form of variance (R²) with linear, gaussian and polynomial kernels. the error value is shown by the linear kernel with an error rate of 35.15%, with a variance value of 64.85%.

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

ABSTRAK

PREDIKSI JUMLAH PENYAKIT DEMAM BERDARAH MENGUNAKAN METODE *SUPPORT VECTOR MACHINE* (STUDI KASUS PENYEBARAN DBD DI SINGAPURA)

Oleh

Danu Sasmita

Demam Berdarah *Dengue* (DBD) adalah penyakit menular yang disebabkan oleh infeksi virus dengue dan ditularkan melalui gigitan nyamuk betina spesies *Aedes aegypti* dan *Aedes albopictus*. Faktor lingkungan menjadi salah satu penyebab tingginya prevalensi penyakit DBD, antara lain tata letak bangunan, tempat penampungan air, lekukan dalam tanah, suhu dan hal-hal lain yang dapat membantu berlangsungnya siklus hidup nyamuk *Aedes*. Peramalan berupa klasifikasi penyebaran penyakit demam berdarah diperlukan untuk mengetahui seberapa besar potensi persebaran penyakit demam berdarah di suatu wilayah yang diharapkan dapat mempermudah untuk menentukan langkah yang tepat dalam proses penanggulangan penyakit demam berdarah, sehingga dapat meningkatkan derajat kesehatan masyarakat. Tujuan dari penelitian ini adalah untuk memprediksi penyebaran penyakit demam berdarah menggunakan metode *Support Vector Machine* (SVM) dengan data curah hujan di Singapura dari tahun 2014 hingga 2018, data cuaca dan data nyeri, membandingkan penelitian ini dengan penelitian sebelumnya oleh Adeline Ong tahun 2014 yang berjudul “ Memprediksi Kasus Demam Berdarah di Singapura”, serta mengetahui hasil prediksi sebaran DBD menggunakan metode SVM berupa nilai varian (R^2) dengan *kernel linier*, *gaussian* dan *polinomial*. nilai error ditunjukkan oleh *kernel linear* dengan tingkat error 35,15%, dengan nilai varian 64,85%.

Kata kunci: Demam Berdarah, *Support Vector Machine*, *K-Fold Cross Validation*.

Judul Skripsi : **PREDIKSI JUMLAH PENYAKIT DEMAM BERDARAH MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (STUDI KASUS PENYEBARAN DBD DI SINGAPURA)**

Nama Mahasiswa : **Danu Sasmita**

Nomor Pokok Mahasiswa : **1617051133**

Jurusan : **Ilmu Komputer**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**



MENYETUJUI

1. Pembimbing

Favorisen R. Lumbanraja, Ph.D.
NIP 19830110 200812 1 002

2. Ketua Jurusan Ilmu Komputer

Didik Kurniawan, S.Si., MT.
NIP 19800419 200501 1 004

MENGESAHKAN

1. Tim Penguji

Ketua

: **Favorisen R. Lumbanraja, Ph.D.**



Penguji I
Sekretaris

: **Dr. Ir. Kurnia Muludi, M.S.Sc**



Penguji II
Bukan Pembimbing

: **Astria Hijriani, S.Kom., M.Kom**



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Suropto Dwi Yuwono, S.Si., M.T.

NIP 197407052000031001



Tanggal Lulus Ujian Skripsi : **8 Februari 2022**

PERNYATAAN

Saya yang bertanda tangan di bawah ini, menyatakan bahwa skripsi saya yang berjudul “**Prediksi Jumlah Penyakit Demam Berdarah Menggunakan Metode *Support Vector Machine* (Studi Kasus Penyebaran DBD di Singapura)**” merupakan karya saya sendiri dan bukan karya orang lain. Semua tulisan yang tertuang di skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila dikemudian hari terbukti skripsi saya merupakan hasil penjiplakan atau dibuat orang lain, maka saya bersedia menerima sanksi berupa pencabutan gelar yang saya terima.

Bandar Lampung, 9 Juni 2022



Danu Sasmita
NPM. 1617051133

RIWAYAT HIDUP



Penulis dilahirkan di Bulokarto pada tanggal 29 Juni 1997, sebagai anak pertama dari dua bersaudara dengan ayah yang bernama Sumarno dan Ibu yang bernama Tri Murni. Penulis menyelesaikan Pendidikan formal pada pertama kali di SD N 1 Wates pada tahun 2009. Pendidikan menengah pertama di SMP N 2 Gadingrejo dan lulus pada tahun 2012, dan melanjutkan pendidikan menengah atas di SMK N 1 Gadingrejo dan lulus pada tahun 2015.

Tahun 2015 Penulis melanjutkan pendidikan di Politeknik Negeri Lampung Jurusan Managemen Informatika, namun pada tahun 2016 penulis melanjutkan pendidikan di Universitas Lampung melalui jalur PMPAP sehingga dibebaskan biaya perkuliahan selama 8 semester. Selama menjadi mahasiswa penulis aktif diberbagai kegiatan dan meraih prestasi sebaga berikut :

1. Duta GenRe Provinsi Lampung Tahun 2017 (Juara Poster Terbaik)
2. Duta GenRe Universitas Lampung Tahun 2017 (Juara Poster Terbaik)
3. Juara 1 Video Kreatif SEMESTA UIN Raden Intan Lampung 2017
4. Juara Terbaik Vidgram Fertival Budaya Unila 2016

PERSEMBAHAN

Segala puji dan syukur saya ucapkan kepada Allah Subhanahu Wa Ta'ala atas segala rahmat dan Ridho-Nya selalu memberikan keyakinan, kekuatan, kesabaran serta kelancaran kepadaku sehingga skripsi ini dapat diselesaikan.

Aku persembahkan karyaku ini untuk :

Kedua Orangtua tercinta yang tak Lelah untuk selalu berdoa dan berusaha demi kesuksesan saya dan adik saya, kasih sayang, nasihat, serta dukungan yang memotivasi untuk terus berjalan hingga garis finish, demi terwujudnya cita-cita yang telah didambakan selama ini.

Adikku yang aku sayangi, terimakasih atas do'a serta dukungan baik moril dan materil sehingga memicu saya untuk terus berjalan melewati segala rintangan yang ada demi gelar dan pekerjaan yang saya dambakan.

Tak lupa terimakasih juga teman-teman seperjuangan, kakak tingkat serta adik tingkat yang telah memberikan semangat dan dukungan.

*Keluarga Ilmu Komputer 2016, Serta almamater tercinta
UNIVERSITAS LAMPUNG*

MOTTO

“Janganlah kamu bersikap lemah dan janganlah pula kamu bersedih hati, padahal kamulah orang-orang yang paling tinggi derajatnya jika kamu beriman.”

(Q.S. Al-Imran : 139)

“Fokus pada perjalanan, bukan tujuan”

(Greg Anderson)

“Ingin menjadi orang lain adalah menyalakan dirimu.”

(Kurt Cobain)

SANWACANA

Alhamdulillahirabbil'amin, puji syukur kehadiran Allah Subhanahu wa ta'ala yang telah melimpahkan rahmat serta hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Prediksi Jumlah Penyakit Demam Berdarah Menggunakan Metode *Support Vector Machine* (Studi Kasus Penyebaran DBD di Singapura)” dengan lancar. Shalawat serta salam semoga senantiasa tercurahkan kepada Nabi Muhammad *Shallahu'alaihi wa sallam*, keluarga, serta sahabatnya, yang kita nantikan syafa'atnya di *yaumul* akhir kelak.

Selama proses penyusunan skripsi, penulis mendapatkan banyak sekali bimbingan, dukungan, serta motivasi yang tentunya sangat membantu penulis dalam proses penyusunan skripsi ini. Rasa syukur dan ucapan terima kasih sedalam-dalamnya penulis tujukan kepada:

1. Keluarga kecilku, Mamak, Bapak, dan adikku Prima Yuda, serta kerabatku yang tak Lelah memberikan semangat serta dukungan moril maupun materil.
2. Bapak Favorisen R. Lumbanraja, Ph.D., selaku dosen pembimbing utama atas kesediaannya dalam membimbing, memberikan ilmu, nasihat serta arahan yang diberikan dengan sabar dan tulus.
3. Dr. Ir. Kurnia Muludi, M.S.Sc., selaku dosen pembahas I yang telah memberikan masukan, saran dan kritik yang membangun demi perbaikan skripsi yang penulis buat.
4. Ibu Astria Hijriani, S.Kom., M.Kom., selaku dosen pembahas II yang juga telah memberikan masukan, saran dan kritik yang membangun demi perbaikan skripsi.
5. Dr. Eng. Suropto Dwi Yuwono, M.T. selaku Dekan FMIPA Universitas Lampung.

6. Bapak Didik Kurniawan, S.Si., MT., selaku Ketua Jurusan Ilmu Komputer FMIPA Universitas Lampung.
7. Dr. rer. nat. Akmal Junaidi, M.Sc., selaku Sekretaris Jurusan Ilmu Komputer FMIPA Universitas Lampung.
8. Bapak dan Ibu Dosen Jurusan Ilmu Komputer Universitas Lampung yang dengan tulus telah memberikan ilmu dan arahan selama penulis menjadi mahasiswa.
9. Ibu Ade Nora Maela, Pak Zainudin, dan Mas Nofal yang telah memudahkan segala urusan administrasi penulis di Jurusan Ilmu Komputer.
10. Teman-teman satu kelompok bimbingan Fanni, Jojo, Rachel, Emes, Friska, Silfia, dan Indah yang tidak ragu untuk saling membantu dalam segala kesulitan pada proses skripsi yang dijalani.
11. Teman organisasi kampus, serta teman satu Angkatan Jurusan Ilmu Komputer 2016 atas kebersamaan, kenangan, dan kerjasama selama menjalankan masa studi di Universitas Lampung.

Penulis menyadari bahwa dalam penulisan skripsi ini masih banyak terdapat kekurangan karena masih terbatasnya pengetahuan, pengalaman, dan kemampuan penulis. Oleh karena itu, penulis sangat mengharapkan saran dan kritik yang membangun untuk penelitian ini sebagai bahan pertimbangan untuk karya tulis yang akan datang. Semoga skripsi ini dapat bermanfaat bagi semua pihak.

Bandar Lampung, 9 Juni 2022

Danu Sasmita
NPM. 1617051133

DAFTAR ISI

Halaman

DAFTAR ISI	xiii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvi
DAFTAR KODE PROGRAM	xvii
I. PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah	4
1.4. Tujuan Penelitian	4
1.5. Manfaat Penelitian	5
II. TINJAUAN PUSTAKA	6
2.1. Penelitian Terdahulu	6
2.2. <i>Forecasting</i>	8
2.3. <i>Demam Berdarah Dengue (DBD)</i>	8
2.5. <i>Machine Learning</i>	14
2.6. <i>Koefisiensi Determinansi (R^2)</i>	16
2.7. <i>Support Vector Machine (SVM)</i>	16
2.7. <i>Cross Validation</i>	18
2.8. Evaluasi	19
III. METODE PENELITIAN	20
3.1. Tempat dan Waktu Penelitian	20
3.3. Metode	25
IV. HASIL DAN PEMBAHASAN	28

4.1 Percobaan	28
4.2 Pembahasan.....	34
V. SIMPULAN DAN SARAN	37
5.1 Simpulan	37
5.2 Saran.....	37
DAFTAR PUSTAKA	35

DAFTAR GAMBAR

Gambar	Halaman
1. Kasus Demam Berdarah di Negara Singapura	2
2. Siklus Hidup Nyamuk <i>Aedes Aegypti</i>	12
3. Tempat Perindukan <i>Aedes Aegypti</i>	14
4. Ilustrasi SVM dalam mencari <i>hyperplane</i>	17
5. Tahapan Metode Penelitian.....	18
6. Ilustrasi Cara Kerja <i>K-Fold Cross Validation</i>	26
7. Grafik Tingkat Kesalahan Metode Prediksi	35

DAFTAR TABEL

Tabel	Halaman
1. Hasil Penelitian Tanpa <i>Feature Selection</i>	7
2. Hasil Penelitian Dengan <i>Feature Selection</i>	7
3. Urutan Klasifikasi Dari Nyamuk <i>Aedes Aegypti</i>	11
4. Jadwal Penelitian.....	21
5. Deskripsi Atribut Penelitian.....	23
6. Contoh Data	234
7. Hasil Menggunakan SVM Dengan <i>Kernel Linear</i>	32
8. Hasil Menggunakan SVM Dengan <i>Kernel Gaussian</i>	32
9. Hasil Menggunakan SVM Dengan <i>Kernel Polynomial</i>	33
10. Lama waktu percobaan	36

DAFTAR KODE PROGRAM

Kode Program	Halaman
1. Memasukkan Data Ke R Studio.....	28
2. Kode Mengubah Nama Set Data.....	29
3. Kode Mengubah Nama Set Data.....	29
4. Pembagian Set Data Dengan 10 <i>Fold Cross Validation</i>	30
5. Membuat Model Regresi SVM <i>Kernel Linear</i>	30
6. Membuat Model Regresi SVM <i>Kernel Gaussian</i>	30
7. Membuat Model Regresi SVM <i>Kernel Polynomial</i>	31
8. Prediksi Dengan Model Regresi.....	31
9. Menemukan Nilai Varian.....	31

I. PENDAHULUAN

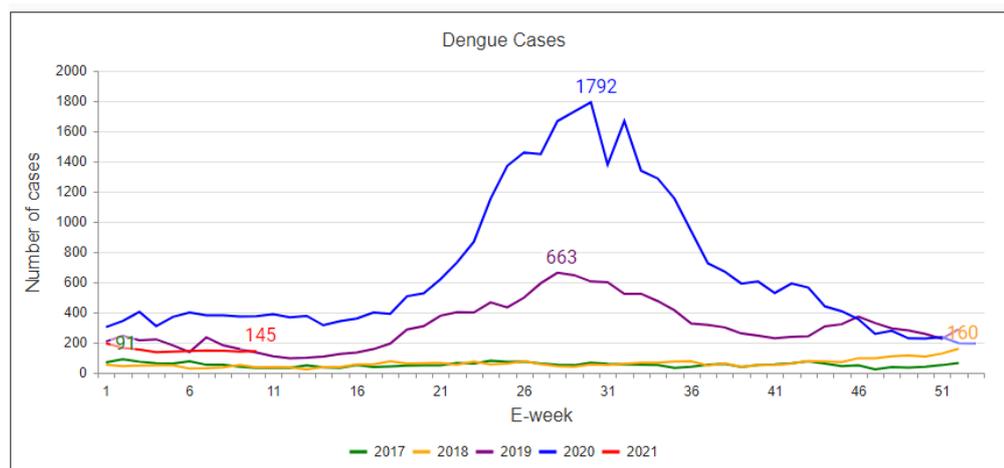
1.1. Latar Belakang

Demam Berdarah Dengue (DBD) merupakan salah satu penyakit menular yang disebabkan oleh salah satu infeksi virus *dengue*. Virus ini ditularkan melalui gigitan nyamuk betina spesies *Aedes aegypti* dan *Aedes albopictus* (Sukohar, 2014). 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. (Candra, 2010).

DBD masih menjadi salah satu masalah penyakit kesehatan di dunia. 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. 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. (WHO, 2017).

DBD ialah penyakit yang disebabkan oleh nyamuk betina berjenis *aedes* yang terpapar virus *dengue*, proses penularan virus *dengue* dari nyamuk *aedes* betina terjadi pada saat mulut nyamuk menusuk bagian kulit manusia sehingga menyebabkan virus bertransmisi (Manuaba et al., 2010). 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 (Sukohar, 2014).

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 (NEA, 2021). Berikut Gambar 1 menunjukkan kasus demam berdarah di Negara Singapura dari Minggu pertama tahun 2017 hingga minggu ke 10 tahun 2020 (NEA, 2021).



Gambar 1. Kasus Demam Berdarah di Negara Singapura (NEA, 2007).

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 (NEA, 2021).

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 (Prasetyani, 2015).

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 (Nugroho et al., 2011).

Tahun 2019, penelitian mengenai persebaran penyakit DBD menggunakan metode SVM pernah dilakukan oleh RM Sulaiman Sani pada penelitian skripsinya, yaitu “Prediksi Jumlah Kasus DBD di Kota Bandar Lampung dengan Menggunakan Metode SVM (*Support Vector Machine*)”. Penelitian itu mengujikan 3 *kernel* berbeda dan masing-masing *kernel* melakukan empat kali pengujian dengan kombinasi jumlah variable yang berbeda setiap percobaannya.

1.2. Rumusan Masalah

Rumusan masalah dalam penelitian ini, yaitu bagaimana melakukan prediksi persebaran penyakit DBD di Singapura dengan data curah hujan, cuaca, dan keluhan penderita DBD.

1.3. Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut :

1. Data yang digunakan merupakan data penyakit demam berdarah di Singapura dari tahun 2014 hingga 2018, beserta data cuaca dan data keluhan penderita yang didapat dari laman <https://towardsdatascience.com/predicting-dengue-in-singapore-f9be0a761ce4>.
2. Metode yang digunakan untuk prediksi yaitu metode SVM, menggunakan *kernel linear*, *gaussian* dan *polynomial*.

1.4. Tujuan Penelitian

Adapun tujuan yang ingin dicapai dalam penelitian ini, yaitu :

1. Melakukan prediksi persebaran penyakit DBD menggunakan metode SVM dengan data curah hujan di Singapura tahun 2014 hingga 2018, data cuaca dan data kesakitan.
2. Membandingkan penelitian ini dengan penelitian sebelumnya oleh Adeline Ong pada tahun 2014 yang berjudul “*Predicting Dengue Cases in Singapore*”.
3. Mengetahui hasil prediksi persebaran penyakit DBD menggunakan metode SVM berupa nilai *varian* (R^2) dengan *kernel linear*, *gaussian* dan *polynomial*.

1.5. Manfaat Penelitian

Manfaat dari penelitian ini antara lain, yaitu:

1. Dapat memberikan model prediksi yang dapat digunakan sebagai referensi untuk melakukan memprediksi persebaran penyakit DBD atau penelitian serupa.
2. Mengetahui hasil prediksi persebaran penyakit DBD menggunakan metode SVM .

II. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Dalam melakukan suatu penelitian diperlukan adanya topik hasil penelitian terdahulu yang berkaitan dengan penelitian ini guna membandingkan hasil varian penelitian tersebut. Berikut beberapa penelitian terdahulu yang menjadi pendukung untuk penelitian ini.

Penelitian sebelumnya yang pertama dilakukan oleh RM Sulaiman Sani (2019) mengenai “Prediksi Persebaran DBD di Kota Bandar Lampung dengan Menggunakan Metode SVM (*Support Vector Machine*)”. Penelitian ini menggunakan tiga *kernel*, yaitu *linear*, *gaussian* dan *polynomial*, dengan jumlah variable setiap percobaannya ialah 10 variabel dan 46 variabel pada percobaan tanpa *feature selection*, 8 variabel dan 44 variabel pada percobaan dengan *feature selection*. Adapun variabel yang digunakan dalam penelitian tersebut meliputi kelembaban_udara_rata_rata, tempmax, tempmin, temperature_rata_rata, kecepatan_angin, curah_hujan, udaramin, udaramax, anginmin, anginmax, luas_area, longitude, latitude, hujanmin, hujanmax, dan jumlah_penduduk. Data yang di gunakan pada penelitian ini adalah data penyakit demam berdarah di kota Bandar Lampung dari tahun 2010 sampai dengan tahun 2018 yang diambil dari Dinas kesehatan Kota Bandar Lampung. Hasil pengujian yang diperoleh dari penelitian tersebut dapat dilihat pada Tabel 1 dan Tabel 2.

Tabel 1. Hasil Penelitian Tanpa *Feature Selection*
(Sani, Lumbanraja, Kurniawan, & Irawati, 2019)

No	Kernel	Tanpa <i>Feature Selection</i>	
		10 var	46 var
1	Linear	70,00%	74,40%
2	Gaussian	70,79%	75,44%
3	Polynomial	70,61%	75,10%

Tabel 2. Hasil Penelitian Dengan *Feature Selection*
(Sani, Lumbanraja, Kurniawan, & Irawati, 2019)

No	Kernel	Dengan <i>Feature Selection</i>	
		8 var	44 var
1	Linear	74,12%	74,61%
2	Gaussian	74,62%	75,51%
3	Polynomial	74,58%	75,15%

Berdasarkan Tabel 1 dan Tabel 2 dapat dilihat bahwa penelitian dengan *feature selection* memiliki hasil nilai varian yang lebih tinggi dari hasil nilai varian penelitian tanpa *feature selection*.

Penelitian kedua yang dilakukan oleh mengenai Adeline Ong (2019) yang meneliti tentang “*Predicting Dengue Cases in Singapore*”. Penelitian tersebut menggunakan data dari tahun 2014 hingga tahun 2018 mengenai kasus demam berdarah, informasi cuaca dan data minat penulisan yang diperoleh dari website. Data minat penelusuran diambil secara mingguan melalui *Google Trends*, meliputi nyeri, demam berdarah, demam, sakit kepala, muntah, mual, ruam, dan sakit mata. Untuk pemilihan model dalam penelitian tersebut menggunakan *5 k-fold cross-validation*. Kesimpulan dari penelitian tersebut dapat menjelaskan bahwa varian dalam log kasus demam berdarah mingguan di Singapura sebanyak 77% dapat digeneralisasikan dengan baik. Serta informasi mengenai minat penelusuran tentang demam berdarah penting digunakan saat melakukan prediksi kasus demam berdarah selanjutnya.

2.2. Forecasting

Forecasting (peramalan) adalah aktivitas untuk melihat kiranya hal apa yang akan terjadi pada masa mendatang terhadap suatu kejadian tertentu, dan merupakan langkah pertama pada proses pengambilan keputusan. Untuk melihat apakah hasil dari peramalan baik atau tidak, dapat digunakan beberapa metode untuk menghitung nilai *forecast error* (kesalahan meramal), semakin kecil *forecast error* maka semakin baik hasil dari ramalan tersebut. Perbedaan jenis ramalan juga dapat menentukan waktu yang digunakan dalam peramalan, apakah termasuk kedalam peramalan jangka pendek, menengah atau panjang. Berdasarkan sifatnya terdapat peramalan yang sangat bergantung pada analisisnya (*kualitatif*) dan ada peramalan yang sangat bergantung pada metodenya (*kuantitatif*) (Iswahyudi, 2016).

2.3. Demam Berdarah Dengue (DBD)

2.3.1 Definisi

Demam Berdarah *Dengue* (*dengue haemorrhagic fever*) atau biasa disingkat DBD adalah penyakit infeksi yang disebabkan oleh virus *dengue* dengan manifestasi klinis adanya demam, nyeri otot, dan/atau nyeri sendi yang disertai dengan gejala leukopenia ruam, limfadenopati, trombositopenia dan diatesis hemoragik. DBD terjadi perembesan plasma yang ditandai dengan hemokonsentrasi (peningkatan hematokrit) atau penumpukan cairan di rongga tubuh (Suhendro et al., 2014).

2.3.2 Etiologi

DBD adalah penyakit infeksi yang disebabkan oleh virus *dengue* yang merupakan bagian dari famili *Flaviviridae*. Ada empat serotipe yang biasanya menyebabkan DBD yaitu DEN-1, DEN-2, DEN-3, dan DEN-4.

Virus *dengue* menginfeksi manusia melalui gigitan vektor perantara nyamuk terutama nyamuk betina dari spesies *Aedes aegypti* dan *Aedes albopictus*. Infeksi pada manusia oleh salah satu serotipe menghasilkan imunitas sepanjang hidup terhadap infeksi ulang oleh serotipe yang sama, tetapi hanya menjadi perlindungan sementara dan parsial terhadap serotipe yang lain (WHO, 2017).

2.3.3 Patogenesis

Patogenesis terjadinya DBD hingga saat ini masih diperdebatkan. Berdasarkan data yang ada, terdapat bukti yang kuat bahwa mekanisme imunopatologis berperan dalam terjadinya DBD dan sindrom renjatan dengue.

Respon imun yang diketahui berperan dalam patogenesis DBD adalah:

- Respon humoral berupa pembentukan antibodi yang berperan dalam proses netralisasi virus, *sitolisis* yang dimediasi komplemen dan sitotoksitas yang dimediasi antibodi. Antibodi terhadap virus *dengue* berperan dalam mempercepat replikasi virus pada *monosit* atau *makrofag*. Hipotesis ini disebut anti-body dependent enhancement (ADE).
- *Limfosit T* baik *T-helper* (CD4) dan *T-sitotoksik* (CD8) berperan dalam respon imun seluler terhadap virus *dengue*. Diferensiasi *T-helper* yaitu Th1 akan memproduksi interferon gamma, IL-2 dan *limfokin*, sedangkan Th2 memproduksi IL-4, IL-5, IL-6, IL-10.
- *Monosit* dan *makrofag* berperan dalam fagositosis virus dengan opsonisasi antibodi. Namun proses fagositosis ini meningkatkan replikasi virus dan sekresi sitokin oleh *makrofag*.

Aktivasi komplemen oleh kompleks imun menyebabkan terbentuknya C3a dan C5a (Suhendro, 2014).

Berdasarkan berbagai data epidemiologi, ada 2 hipotesis yang menjelaskan tentang patogenesis DBD. Kedua teori tersebut adalah “the secondary

heterotypic antibody dependent enhancement of a dengue virus infection” dan gabungan efek jumlah virus, virulensi virus, dan respons imun inang (Achmadi et al., 2010).

Virus *dengue* masuk ke dalam tubuh inang kemudian mencapai sel target yaitu *makrofag*. Sebelum mencapai sel target maka respon imun non-spesifik dan spesifik tubuh akan berusaha menghalanginya. Aktivitas komplemen pada infeksi virus dengue diketahui meningkat seperti C3a dan C5a mediator-mediator ini menyebabkan terjadinya kenaikan permeabilitas kapiler celah endotel melebar lagi. Akibat kejadian ini maka terjadi ekstrasvasi cairan dari *intravaskuler* ke *ekstravaskuler* dan menyebabkan terjadinya tanda kebocoran plasma seperti *hemokonsentrasi*, *hipoproteinemia*, *efusi pleura*, *asites*, penebalan dinding *vesica felea* dan *syok hipovolemik*. Kenaikan permeabilitas kapiler ini berimbas pada terjadinya hemokonsentrasi, tekanan nadi menurun dan tanda syok lainnya merupakan salah satu patofisiologi yang terjadi pada DBD (Achmadi et al., 2010).

2.3.4 Manifestasi Klinik

Manifestasi klinik infeksi virus *dengue* dapat bersifat *asimtomatik*, atau dapat berupa demam ada demam yang tidak khas, demam *dengue*, demam berdarah *dengue*, bahkan sindrom syok *dengue* (SSD) dan sindrom *dengue* diperluas. Pada umumnya orang yang terkena infeksi dari virus *dengue* ini akan mengalami fase demam selama 2-7 hari. Akhir fase demam adalah periode fase kritis pada DBD pada saat fase demam mulai menurun dan orang yang terkena infeksi virus DBD seakan – akan sembuh, tetapi pada periode masa inilah memiliki resiko tinggi untuk terjadi kejutan jika tidak mendapatkan pengobatan yang ade kuat (Suhendro, 2014).

Pada dasarnya gejala dari infeksi virus *dengue* adalah *asimtomatik*. Masa inkubasi dari virus dengue ini sekitar 4-7 hari tapi bisa di kisaran 14 hari.

Gejala dan tanda dari infeksi virus *dengue* ini adalah demam mendadak disertai sakit kepala, nyeri retro-orbital, iyalasi mialgia dan artralgia, pembilasan wajah, anoreksia, sakit perut dan mual, serta ruam. Pada pemeriksaan laboratorium biasanya terdapat leukopenia dan trombositopenia (WHO, 2016).

Tanda penting dari DBD adalah meningkatnya permeabilitas vaskular sehingga terjadi kebocoran plasma, volume intravaskular berkurang, dan syok di kasus yang parah. Kebocoran plasma bersifat unik karena plasma yang bocor selektif, yaitu di pleura dan rongga abdomen serta periodonya pendek (24-48 jam). Pemulihan cepat dari syok tanpa sequele dan tidak adanya inflamasi pada pleura dan peritoneum mengindikasikan mekanisme yang terjadi adalah perubahan fungsi integritas vaskular, bukan kerusakan struktural dari endotel (WHO, 2014).

2.4 Nyamuk *Aedes Aegypty*

2.4.1 Taksonomi *Aedes sp*

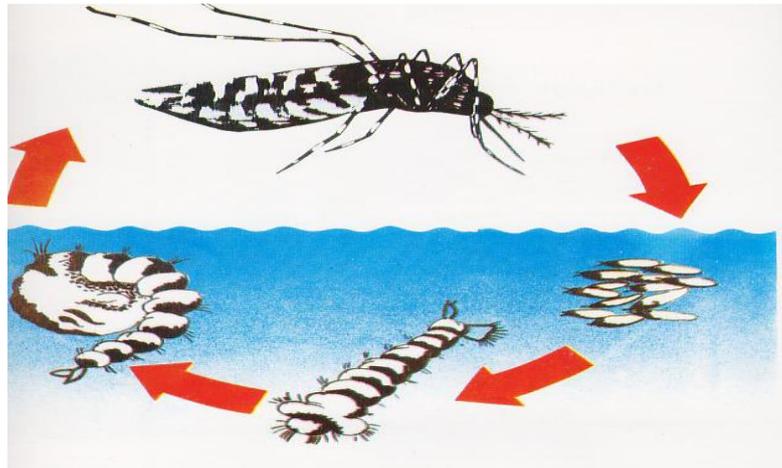
Nyamuk *Aedes sp* tersebar di seluruh dunia dan diperkirakan mencapai 950 spesies. Nyamuk ini dapat menyebabkan gangguan gigitan yang serius terhadap manusia dan binatang, baik di daerah tropik dan daerah beriklim lebih dingin (Inge Susanto et al., 2011). Berikut Tabel 3 urutan klasifikasi dari nyamuk *Aedes Aegypti* adalah sebagai berikut :

Tabel 3. Klasifikasi Nyamuk *Aedes Aegypti* (Inge Susanto et al., 2011)

Kingdom	Animalia
Filum	Arthropoda
Kelas	Insecta
Ordo	Diptera
Famili	Culicidae
Genus	<i>Aedes</i>
Spesies	<i>Aedes Aegypty</i>

2.4.2 Morfologi dan Siklus Nyamuk *Aedes Aegypti*

Masa pertumbuhan dan perkembangan nyamuk *Aedes aegypti* dari telur menjadi nyamuk dewasa dibagi menjadi empat tahap, yaitu telur, jentik, pupa, dan nyamuk dewasa. Nyamuk *Aedes aegypti* termasuk kedalam metamorphosis sempurna (holometabola) (Soedarto,2012). Berikut siklus hidup nyamuk *Aedes aegypti* pada Gambar 2.



Gambar 2. Siklus Hidup Nyamuk *Aedes Aegypti* (Kemenkes RI, 2016).

2.4.7 Tempat Perindukan Nyamuk *Aedes aegypti*

Aedes aegypti suka bertelur di air jernih yang tidak berhubungan langsung dengan tanah. Tempat perkembangbiakan utama ialah tempat-tempat penampungan air bersih yang tertampung di suatu bejana di dalam atau di sekitar rumah atau di tempat-tempat umum yang berjarak tidak melebihi 500 meter dari rumah (Kemenkes RI, 2016).

Keberadaan tempat penampungan air (TPA) di lingkungan rumah sangat berperan dalam kepadatan jentik *Aedes sp*, karena semakin banyak tempat penampungan air (TPA) akan semakin banyak tempat perindukan dan akan semakin padat populasi nyamuk *Aedes sp*. Semakin padat populasi nyamuk *Aedes sp*, maka semakin tinggi pula risiko terinfeksi virus DBD dengan

waktu penyebaran lebih cepat sehingga jumlah kasus penyakit DBD cepat meningkat yang pada akhirnya mengakibatkan Kejadian Luar Biasa (KLB) penyakit DBD (Majid, 2017).

Jumlah TPA dalam penelitian Wisfer menunjukkan bahwa rumah dengan banyak TPA memiliki peluang lebih besar dibandingkan rumah yang memiliki sedikit TPA *terhadap* keberadaan larva karena keberadaan kontainer sangat berperan dalam kepadatan vector nyamuk *Aedes aegypti*, semakin banyak kontainer maka akan semakin banyak tempat perindukan dan akan semakin padat populasi nyamuk *Aedes aegypti* (Wisfer et al., 2014).

Menurut Kemenkes RI (2016), jenis tempat perindukan nyamuk *Aedes aegypti* dapat dikelompokkan sebagai berikut:

- Tempat penampungan air untuk keperluan sehari-hari (TPA Sehari-hari), seperti drum, tangki reservoir, tempayan, bak mandi/WC dan ember.
- Tempat penampungan air yang bukan untuk keperluan sehari-hari (TPA bukan untuk Sehari-hari), seperti tempat minum burung, vas bunga, perangkap semut dan barang-barang bekas (ban, kaleng, botol, plastik dan lain-lain).
- Tempat penampungan air alamiah, seperti lubang pohon, lubang batu, pelepah daun, tempurung kelapa, pelepah pisang dan potongan bambu.

Menurut Ditjen P2PL (2014), tempat perkembangbiakan (*breeding places*) jentik *Aedes aegypti* dibedakan sebagai berikut:

- **Artifisial (Buatan)**
Tempat perkembangbiakan jentik buatan adalah tempat yang dibuat oleh *manusia* dimana dapat menampung air dan jernih yang kemudian digunakan oleh nyamuk *Aedes* untuk berkembangbiak. Adapun contoh tempat perkembangbiakan jentik buatan yakni bak mandi, ember,

dispenser, kulkas, ban bekas, pot/vas bunga, kaleng, plastik, dan lain-lain.

- Natural (Alamiah)

Tempat perkembangbiakan jentik alamiah adalah tempat yang dapat *menampung* air jernih dan telah tersedia di lingkungan pemukiman. Adapun contoh tempat berupa tempat perindukan nyamuk pada tempat alami yakni tanaman yang dapat menampung air, ketiak daun, tempurung kelapa, lubang bambu, ataupun pada pelepah daun. Berikut Gambar 7 contoh tempat perindukan nyamuk *Aedes aegypti*.



Gambar 3. Tempat Perindukan *Aedes Aegypti* (Kemenkes RI, 2016).

2.5. Machine Learning

Machine Learning adalah metode dalam kecerdasan buatan yang digunakan untuk membuat aplikasi yang memiliki sifat seperti manusia untuk menyelesaikan masalah dan melakukan otomasi. Sama seperti manusia, *machine learning* memerlukan proses belajar, dan latihan atau *training* menggunakan data yang biasa disebut dengan *data training* (Ahmad, 2017).

Machine Learning mengenali pola dari caranya mempelajari kasus dari data yang telah ada untuk dianalisa dan menggunakan pola yang ada untuk memecahkan masalah dari suatu pertanyaan-pertanyaan. Program yang dibuat menggunakan metode *machine learning* diantaranya: *Online Fraug Detection*, *Diagnosa Medis*, *Mobil Kendali Otomatis*, *Segmentasi Pelanggan* dan *Mesin Pencarian*. *Machine Learning* dibagi menjadi dua jenis, yaitu *supervised learning* dan *unsupervised learning*.

1. *Supervised Learning*

Supervised learning merupakan cara belajar bagaimana melakukan memprediksi berdasarkan *input* dan *output* yang ada. Tujuannya adalah membangun model yang dapat menghasilkan *output* yang benar untuk data *input*. Contoh algoritma dalam *supervised learning*, yaitu *decision trees*, *neural network*, *bayesian filtering*, dan *support vector machine* (Segaran, 2007). Terdapat dua jenis *supervised learning*, yaitu klasifikasi dan regresi.

Regresi adalah model yang digunakan untuk memprediksi nilai yang kontinu. Contohnya memprediksi harga sebuah rumah dengan mempertimbangkan ukuran dan isi rumah. *Regression* dibagi menjadi lima jenis, yaitu *Simple Linear Regression*, *Polynomial Regression*, *Support Vector Machine*, *Decision Tree Regression*, *Random Forest Regression*. *Classification* adalah hasil yang diberikan bersifat diskret, diberikan kategori dan label pada data yang ada sehingga keputusan yang diambil berdasarkan kategori dan label yang telah ada. Klasifikasi bertujuan untuk memisahkan data menjadi kelas-kelas tertentu (Segaran, 2007).

2. *Unsupervised Learning*

Unsupervised learning merupakan cara untuk mencari pola berdasarkan data yang tidak memiliki jawaban didalamnya. Metode *unsupervised learning* yang paling umum adalah analisa cluster, yang digunakan pada analisa data untuk mencari pola-pola tersembunyi atau pengelompokan dalam data. Tujuan pembelajaran adalah membangun model yang dapat

menemukan variabel tersembunyi pada data pelatihan. Contoh *unsupervised learning* yaitu *K-means*, *Fuzzy clustering*, dan *Hierarchical clustering* (Nurhayati at al., 2019).

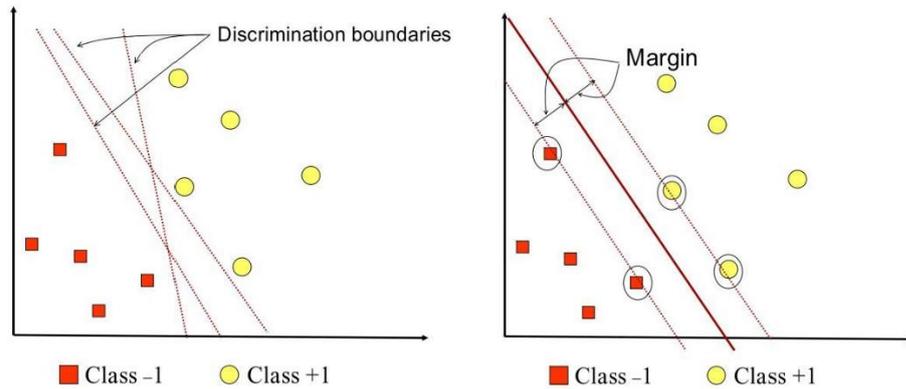
2.6. Koefisiensi Determinansi (R^2)

Koefisiensi Determinansi adalah Media untuk menaksir ukuran suatu variabel yang bersifat bebas atas variabel yang bersifat terikat. Dengan nilai koefisien determinansi yang dimulai dari 0 dan maksimal 1. Variabel independen dapat dikatakan berpengaruh besar jika nilai yang didapat semakin mendekati 1. (Rizal & Alexander, 2019).

2.7. Support Vector Machine (SVM)

Support vector machine adalah *machine learning* baru untuk masalah klasifikasi dua kelompok yang menggabungkan 3 ide, yaitu *the solution technique* dari *optimal hyperlanes* (yang memungkinkan perluasan vektor solusi pada vektor dukungan), gagasan *convolution of the dot-product* (yang memperluas permukaan solusi dari linier ke non- linear), dan gagasan *soft margin* (untuk memungkinkan kesalahan pada set pelatihan) (Vapnik & Cortes, 1995).

SVM melakukan transformasi data dari ruang masukan ke ruang yang lebih tinggi dari segi dimensinya yang disebut ruang vector dan melakukan tindakan optimasi (Nugroho 2003). Ilustrasi SVM dalam usahanya menemukan *hyperplane* terbaik dapat dilihat pada Gambar 4.



Gambar 4. Ilustrasi SVM dalam mencari *hyperplane* (Nugroho et al, 2003).

Dari Ilustrasi gambar 4, *hyperplane* dapat dikatakan menjadi pemisah terbaik karena garis solid yang terdapat pada ilustrasi tersebut berada persis diantara dua class. Hal tersebut didapat dari hasil pengukuran jarak *hyperplane* dengan patem yang terletak paling dekat dari masing-masing kelas dan berusaha menemukan poin maksimalnya (Nugroho et al, 2003).

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 (Parapat et al, 2018). Berikut Persamaan (1) :

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

Serta kelas positif yang diilustrasikan pada kelompok berwarna kuning yang bernilai +1 (Parapat et al, 2018). Juga dapat dilihat pada Persamaan (2) :

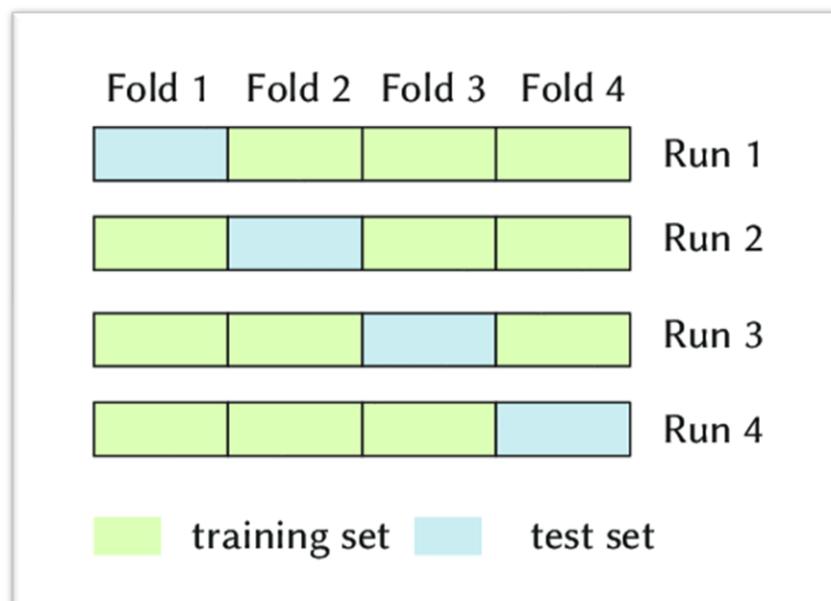
$$X_i \cdot W + b \geq 1 \text{ untuk } Y_i = 1 \dots\dots\dots(2)$$

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*.

2.7. Cross Validation

Cross-validation adalah cara untuk mengevaluasi dan membandingkan algoritma *learning* menggunakan metode statistik dengan membagi data menjadi dua segmen yaitu *set training* dan *validation*. *Set training* digunakan untuk mempelajari atau melatih suatu model, sedangkan *set validation* digunakan untuk memvalidasi model. Dalam *cross-validation*, *set training* dan *validation* harus *cross-over* dalam putaran berturut-turut hingga setiap titik data memiliki peluang untuk divalidasi (Santoso, 2009).

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 (Kohavi & Provost, 1998). Ilustrasi mengenai cara kerja *K-fold cross-validation* dapat dilihat pada Gambar 5.



Gambar 5. Ilustrasi Cara Kerja *K-Fold Cross Validation* (Researchget, 2015).

2.8. Evaluasi

Terdapat beberapa model evaluasi yang digunakan dalam penelitian, untuk mengetahui model yang terbaik dibutuhkan evaluasi dengan mengoperasikan SSE (*Sum of Square Error*) dan Koefisien Determinan (R^2).

SSE digunakan untuk melihat tingkat kesalahan penggunaan model peramalan dengan cara memperkirakan selisih total dari data actual terhadap data hasil prediksi (Sari et al., 2014). Berikut Persamaan (3) :

$$SSE = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \dots\dots\dots(3)$$

Namun sebelumnya perlu diketahui berapa total dari variabilitas dalam data untuk mengukur variabilitas yang sebelumnya tidak dijabarkan dalam model prediksi, dengan menggunakan fungsi SST (Total Sum of Squares) (Lewis-Beck et al., 2012). Yang terlihat pada Persamaan (4).

$$SST = \sum_{i=1}^n (Y_i - \bar{Y})^2 \dots\dots\dots(4)$$

Variabel Y_i merupakan data aktual dan variabel \hat{Y} adalah hasil prediksi. *Variable* yang dapat mempengaruhi *variable* terikat dapat diukur melalui pengujian koefisien determinan yang memiliki nilai mulai dari 0 hingga 1, atau dapat ditulis $0 \leq R^2 \leq 1$. Jika $R^2 = 0$ variabel tidak mempengaruhi *variable* terikat, jika nilai R^2 semakin bertambah hingga kurang dari 1 maka semakin kuat pengaruh *variable* tersebut terhadap *variable* terikat, namun jika nilai R^2 semakin rendah hingga kurang dari 0 maka semakin rendah pengaruh *variable* tersebut terhadap *variable* terikat (Ariska & Indriasari, 2019).

Koefisien Determinan merupakan faktor penting yang menunjukkan seberapa baik sebuah model prediksi mempertimbangkan *variable* variasi dalam *variable* respon (Zhang, 2017). Persamaan (5) merupakan fungsi koefisien determinan. $R^2 = 1 - SSE / SST \dots\dots\dots(5)$

III. METODE PENELITIAN

3.1. Tempat dan Waktu Penelitian

3.1.1 Tempat Penelitian

Penelitian dilakukan di dua tempat, yang pertama di kediaman penulis yang berada di Jalan Pelita 1, Kecamatan Gadingrejo, Kabupaten Pringsewu, Lampung. Dan yang kedua dilakukan di Laboratorium Komputer Rekayasa Perangkat Lunak Jurusan Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung yang beralamatkan di jalan Soemantri Brojonegoro No. 1 Gedung Meneng, Bandar Lampung.

3.1.2 Waktu Penelitian

Waktu penelitian ini di laksanakan pada minggu pertama bulan April 2020 hingga minggu ke empat bulan November 2021. Waktu lebih rinci dapat dilihat pada Tabel 4.

Rincian waktu berdasarkan chart pada Tabel 4 yaitu, studi literatur dan penulisan laporan Bab 1 hingga 3 dilakukan pada awal bulan April hingga minggu ke empat bulan September. Sembari melakukan studi literatur dan penulisan laporan, dilakukan pengumpulan data mulai minggu pertama bulan April hingga minggu ke empat bulan Agustus. Proses Cleaning data Dilakukan pada bulan September, dari minggu pertama hingga minggu keempat bulan Januari 2021. Dilanjutkan seminar usul pada bulan Januari minggu keempat.

Setelah Seminar Usul, dilakukan proses *K-Fold Cross Validation* yang dimulai dari minggu keempat bulan Januari hingga minggu kedua bulan Februari 2021. Membangun model SVM dilakukan selama tiga minggu dimulai dari awal bulan Februari 2021, dilanjutkan dengan evaluasi kinerja pada minggu pertama hingga minggu kedua bulan Maret 2021. Proses penulisan laporan Bab 4 dan 5 dilakukan setelah proses seminar usul yaitu pada minggu kedua bulan Maret hingga minggu ketiga bulan November 2021.

3.2. Data dan Alat

3.2.1 Data

Data dalam penelitian ini didapat dari situs <https://github.com/adelweiss/Dengue/tree/master>. Dimana data tersebut sebelumnya digunakan oleh Adeline Ong pada tahun 2019 dalam penelitiannya yang berjudul "*Predicting Dengue Cases in Singapore*". Data yang akan diolah dalam penelitian ini disimpan di dalam 15 variabel, Berikut deskripsi mengenai atribut penelitian yang digunakan terdapat pada Tabel 5.

Tabel 5. Deskripsi Atribut Penelitian

Nama Atribut	Deskripsi
<i>dengue_cases</i>	Data kasus demam
<i>Rainfall</i>	Kumpulan data curah hujan
<i>mean_temp</i>	Nilai rata-rata dari temperature disimpan disini
<i>max_temp</i>	Nilai temperature maksimal disimpan disini
<i>min_temp</i>	Nilai temperature minimal
<i>Fever</i>	Data kasus demam berdarah
<i>Nausea</i>	Angka kejadian mual
<i>Headache</i>	Data kejadian sakit kepala
<i>ache_pain</i>	Data kasus nyeri
<i>eye_pain</i>	angka kejadian sakit mata
<i>Dengue</i>	angka kejadian demam
<i>Rashes</i>	data kejadian ruam pada kulit
<i>Vomiting</i>	kasus kejadian muntah
<i>Dangue_cases_+1</i>	kasus demam berdarah ditambah satu
<i>temp_range</i>	data jarak temperature

Berikut sampel data penelitian yang meliputi 15 variabel dapat dilihat dalam Tabel 6.

Tabel 6. Contoh Data

Dengue_case	rainfall	Mean_ temp	Max_ temp	Min_ temp	fever	nausea	headache	Ache_ pain	Eye_pain	dengue	rashes	vomiting	Dengue_ cases_+1	Temp_range
127	1.628571	27.48571	31.35714	25.44286	77	77	76	84	61	28	60	96	160	5.914286
107	7	27.45714	30.84286	24.88571	67	92	56	87	81	26	58	86	127	5.957143
113	7.314286	27.15714	31.12857	23.95714	82	70	70	77	59	29	54	89	107	7.171429
109	5.914286	27.22857	31.11429	24.34286	76	64	80	82	76	27	62	78	113	6.771429
97	3.8	27.45714	31.6	25.01429	67	56	70	82	72	20	67	73	109	6.585714
78	2.057143	27.34286	31.17143	25.31429	73	80	70	89	62	22	51	81	97	5.857143
71	0.628571	28.11429	32.21429	25.77143	74	75	69	87	48	17	44	65	78	6.442857
74	4.428571	27.52857	31.11429	24.21429	67	91	74	81	23	23	50	77	71	6.9
78	26.31429	26.82857	30.91429	23.92857	68	78	68	94	61	21	53	85	74	6.985714
60	1.257143	29.47143	33.34286	27.1	65	74	72	89	39	19	45	75	78	6.242857
51	0	29.07143	32.54286	26.55714	57	90	74	87	42	15	56	64	60	5.985714

3.2.2 Alat

Alat yang digunakan dalam penelitian ini meliputi hardware dan software sebagai berikut:

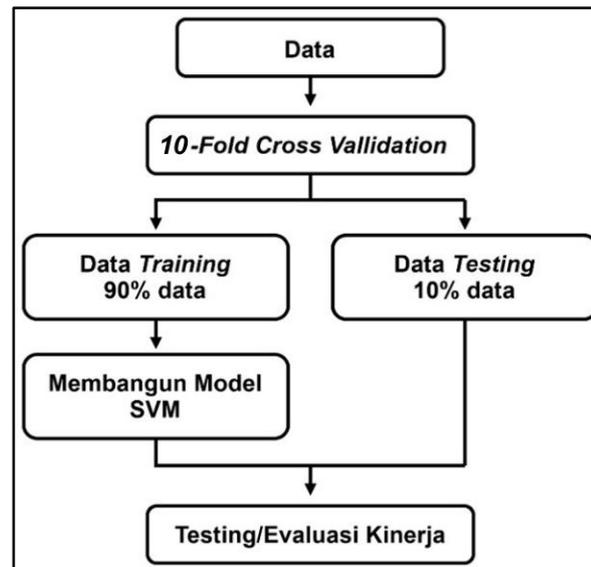
- Hardware 1 : Prosesor Intel® Core™ i3-6006U CPU @ 2.20GHz, Core Speed 798.10 MHz, RAM 8.00 GB DDR3, HDD 500 GB, Layar 14 inch, Core HD Graphic 3000.
- Hardware 2 : Prosesor Intel® Core™ i3-6006U, RAM 8.00 GB, HDD 500 GB.
- Software : Sistem operasi Windows 10 Profesional, Sistem Operasi Windowsn Server 2007, R studio 1.4.1106, Microsoft Office Excel 2019, R versi 3.5.1.

Dalam menggunakan R versi 3.5.1. digunakan beberapa *package* yang memiliki fungsi masing-masing dalam proses penelitian. Berikut *package* yang digunakan:

- *Package readr 1.4.0*, digunakan untuk membaca file dengan format csv.
- *Package caret 6.0-86*, digunakan untuk melatih model prediktif “regresi”.
- *Package ggplot 2 3.3.3*, digunakan untuk melakukan visualisasi terhadap data.

3.3. Metode

Metode yang dilakukan dalam melakukan penelitian tentang prediksi persebaran penyakit demam berdarah di Singapura menggunakan metode SVM dilakukan dengan beberapa tahap yang ditunjukkan pada Gambar 6, sebagai berikut:



Gambar 6. Tahapan Metode SVM dengan *10-Fold Cross Validation*.

Penjelasan dari tahapan metode penelitian pada Gambar 5 adalah sebagai berikut:

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. *K-Fold Cross Validation*

Pada proses ini data dibagi menjadi dua bagian untuk proses *testing* dan *training*. Pada proses *training* data yang digunakan sebanyak 90%, sedangkan pada proses *testing* data yang digunakan hanya sebanyak 10%.

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.

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*.

V. SIMPULAN DAN SARAN

5.1 Simpulan

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.

5.2 Saran

Dari penelitian yang telah dilakukan terdapat saran untuk meningkatkan penelitian selanjutnya sebagai berikut :

1. Waktu percobaan paling lama dalam penelitian ini mencapai 10 hari dengan komputer yang harus hidup tanpa henti selama percobaan. Disarankan penelitian selanjutnya dapat menggunakan perangkat dengan spesifikasi yang lebih tinggi untuk mempercepat waktu percobaan
2. Dapat menggunakan *feature selection* untuk memperbaiki nilai varian.
3. Disarankan untuk penelitian selanjutnya dapat menggunakan metode lain seperti *deep learning* yang diharapkan dapat meningkatkan nilai varian pada percobaan.

DAFTAR PUSTAKA

- Ahmad, A. (2017). Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning. *Jurnal Teknologi Indonesia*, October, 3.
- Abd. Majid HR. Lagu, Damayati, D. S., & Wardiman, M. (2017). Hubungan jumlah penghuni, jumlah tempat penampungan air dan pelaksanaan 3M plus dengan keberadaan jentik nyamuk *Aedes sp* di Kelurahan Balleangin Kecamatan Balocci Kabupaten Pangkep. *Higiene: Jurnal Kesehatan Lingkungan*, 3(1), 22-29.
- Achmadi, F. U., Sudjana, P., Sukowati, S., Wahyono, M. Y., Haryanto, B., Mulyono, S. (2010). *Buletin Jendela Epidemiologi* vol: 2:30-1. Kemenkes RI : Jakarta.
- Adeline Ong (2019). Predicting Dengue Case in Singapore. <https://towardsdatascience.com/predicting-dengue-in-singapore-f9be0a761ce4>
- Apps.who.int. Dengue and severe dengue. (2019, April). Diakses pada 14 Februari 2022, dari <https://apps.who.int/mediacentre/factsheets/fs117/en/index.html>
- Ariska, Y. I., & Indriasari, N. (2019). Analisis Pengaruh Penggunaan Sistem Informasi Manajemen Daerah (Simda) Terhadap Kualitas Laporan Keuangan Pemerintah Daerah (Studi Kasus Dinas Pekerjaan Umum dan Penataan Ruang Provinsi Bengkulu). *Managament Insight: Jurnal Ilmiah Manajemen*, 12(2), 51–63. <https://doi.org/10.33369/insight.12.2.51-63>
- Candra, A. (2010). Demam Berdarah Dengue : Epidemiologi, Patogenesis, dan Faktor Risiko Penularan, 2(2), 110–119.
- Iswahyudi, C. (2016). *Pengantar Forecasting (Teknik Peramalan)*.
- Inge Susanto, Is Suhariah Ismid, Pudji K. Sjarifuddin, Saleha Sungkar. 2011. *Parasitologi Kedokteran*, Edisi Keempat. Jakarta: Fakultas Kedokteran UI, Jakarta.
- Indri Monika Parapat, Muhammad Tanzil Furqon, and Sutrisno, “Penerapan Metode Support Vector Machine (SVM) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer.*, vol. 2, no. 10, 3163–3169, 2018.
- Kohavi, R., & Provost, F. (1998). Glossary of terms: Special Issue on Applications of Machine Learning and the Knowledge Discovery Process. *Machine Learning*, 271-274.
- Manuaba, I. D., Sutirtayasa, I. W. P., & Dewi, D. R. (2010). *Immunopatogenesis infeksi virus dengue*. 1–12.

- Nea.goc.sg. Dengue Cases. Diakses pada 15 Juni 2020, dari <https://www.nea.gov.sg/dengue-zika/dengue/dengue-cases>
- Nugroho, A. S., Witarto, A. B., & Handoko, D. (2011). Support Vector Machine Teori dan Aplikasinya dalam Bioinformatika1. *Proceedings of the 2011 Chinese Control and Decision Conference, CCDC 2011*, 842–847. <https://doi.org/10.1109/CCDC.2011.5968300>
- Nurhayati, Busman, and R. P. Iswara, “Pengembangan Algoritma Unsupervised Learning Technique Pada Big Data Analysis di Media Sosial sebagai media promosi Online Bagi Masyarakat,” *J. Tek. Inform.*, vol. 12, no. 1, pp. 79–96, 2019.
- Prasetyani, R. D. (2015). Faktor-Faktor Yang Berhubungan Dengan Kejadian Demam Berdarah Dengue. *Journal Majority*, 4(7), 61–66. <https://juke.kedokteran.unila.ac.id>
- Pusdatin Kemenkes RI. (2017). *InfoDatin Demam Berdarah Dengue 2017*. <http://www.pusdatin.kemkes.go.id/download.php?file=download/pusdatin/infodatin/InfoDatin-Situasi-Demam-Berdarah-Dengue.pdf>
- Taufiq Risal, & Austin Alexander. (2019). Pengaruh Persepsi Bagi Hasil, Promosi dan Kualitas Pelayanan Terhadap Minat Penggunaan Jasa Perbankan Syariah Tabungan Mudharabah Pada Mahasiswa Universitas Potensi Utama. *Jurnal Samudra Ekonomika*, 3(2), 118-130. <https://doi.org/10.0123/jse.v3i2.1886>
- Santoso, B. (2007). *Data Mining, Teori dan Aplikasi*. Jakarta: Graha Ilmu.
- Sari, I. P., Andrizal, & Devianto, D. (2014). Perancangan Sistem Pengenalan Wajah Manusia Menggunakan Web Camera dengan Metode Summary Squared Error (SSE). *Sistem Komputer*, 1–8.
- Sukohar, A. (2014). Fakultas Kedokteran Universitas Lampung Demam Berdarah Dengue (DBD) Fakultas Kedokteran Universitas Lampung. *Medula*, 2(2), 1–15.
- Sulaiman Sani, R. (2019). PREDIKSI PERSEBARAN DEMAM BERDARAH DI KOTA BANDAR LAMPUNG DENGAN METODE SUPPORT VECTOR MACHINE. *Ayan*, 8(5), 55.
- Soedarto. 2011. *Buku ajar Parasitologi kedokteran*. Jakarta: Sagung Seto
- Suhendro, Nainggolan, L., Chen, K., & Pohan, H. T. (2014). *Buku Ajar Ilmu Penyakit Dalam Edisi IV*. Jakarta
- T. Segaran. (2007). *Programming Collective, Intelligence: Building Smart Web 2.0 Applications*. O'Reilly Media, Inc.
- Vapnik, V., & Cortes, C. (1995). Present and Future Energy Scenario in India. *Journal of The Institution of Engineers (India): Series B*, 95(3), 247–254. <https://doi.org/10.1007/s40031-014-0099-7>

Wisfer, Ibrahim E, Selomo M (2014). Hubungan Jumlah Penghuni, Tempat Penampungan Air Keluarga dengan Keberadaan Larva *Aedes aegypti* di wilayah Endemis DBD Kota Makassar.

Zhang, D. (2017). A Coefficient of Determination for Generalized Linear Models. *American Statistician*, 71(4), 310–316.
<https://doi.org/10.1080/00031305.2016.1256839>