

IMPLEMENTASI METODE *SUPPORT VECTOR MACHINE* DALAM PREDIKSI PERSEBARAN DEMAM BERDARAH DI KOTA BANDAR LAMPUNG

¹Favorisen R. Lumbanraja, ²RM Sulaiman Sani, ³Didik Kurniawan, ⁴Anie Rose Irawati.

Jurusan Ilmu Komputer, FMIPA,

Universitas Lampung, Bandar Lampung, 35141

¹favorisen.lumbanraja@fmipa.unila.ac.id, ²rmsulaimansani@gmail.com, ³didik.kurniawan@fmipa.unila.ac.id,
⁴anie.roseirawati@fmipa.unila.ac.id

Abstract.

Dengue fever is a one of the dangerous diseases and very often causes casualties every year, especially in the tropics or subtropics countries. Dengue fever cases increase during the rainy season, many factors affect the spread of dengue fever, such as vegetation, population and landfills. The aim of this research is to predict the number of cases of Dengue fever using support vector machine. The data used are dengue data in Bandar Lampung City, weather data, population data and distance matrix data between dengue fever events with each other. The amount of data used is 1,080 data with 3 kernels: linear, gaussian and polynomial. In this study four experiments were carried out, the first two experiments were carried out without Feature Selection and the next two experiments were carried out with Feature Selection. After the experiment was found The best performance in the experiment with Feature Selection with 44 Variables. From the experiments conducted, Gaussian kernel achieved the highest R^2 value which is 75.52%, while the Linear kernel and Polynomial Kernel achieved R^2 value of 74.61% and 75.15%, respectively.

Keywords: *Machine Learning, Prediction of Dengue Fever, Regresi, Support Vector Machine.*

1 PENDAHULUAN

Demam berdarah merupakan suatu penyakit yang berbahaya dan selalu memakan korban setiap tahunnya terutama di daerah tropis ataupun subtropis. Peningkatan penyakit demam berdarah terjadi selama musim hujan. Namun, banyak faktor-faktor daerah yang mempengaruhi penyebaran penyakit DBD, termasuk faktor-faktor seperti vegetasi, tipe perumahan, kepadatan penduduk, fasilitas pembuangan sampah, dan sebagainya [1].

Data dari Dinas Kesehatan Kota Bandar Lampung menyebutkan pada 2014, jumlah pasien demam berdarah di Bandar Lampung mencapai 763 orang dan 16 orang meninggal. Pada 2015, jumlah demam berdarah, pasien demam di Bandar Lampung mencapai 413 orang dan 7 orang meninggal. Pada tahun 2016, peningkatan jumlah pasien demam berdarah di Bandar Lampung mencapai 1.111 orang dan 11 orang yang meninggal, jumlah tersebut merupakan yang tertinggi dibandingkan dengan kabupaten lain [2].

Forecasting (peramalan) merupakan cara untuk meramalkan kejadian masa depan berdasarkan data masa lalu yang diukur secara *periodic* dan akan membentuk suatu rangkai waktu data atau yang biasa disebut dengan *Time Series*. Rangkai waktu merupakan suatu pengamatan terhadap variabel tunggal yang diukur secara teratur selama periode waktu tertentu.[3].

Salah satu metode yang dapat digunakan dalam proses prediksi adalah *Support Vector Machine*. Metode ini merupakan teknik berbasis *machine learning* yang masih jarang digunakan untuk melakukan proses prediksi, baik digunakan dalam kasus klasifikasi maupun regresi, dan sangat populer belakangan ini. *Support Vector Machine* merupakan satu dari sekian banyak metode yang dapat digunakan untuk menyelesaikan berbagai macam permasalahan diantaranya ialah masalah prediksi. Dalam menyelesaikan sebuah permasalahan, *Support Vector Machine* mampu menangani permasalahan *non-*

linier dengan adanya fungsi *kernel* yang membuat metode ini dapat digunakan untuk peramalan *time series* [4].

Penelitian yang sama pernah dilakukan oleh Harold Situmorang pada tahun 2015 dalam Tesisnya yang berjudul “Klasifikasi Wilayah Epidemis Demam Berdarah Menggunakan Metode *Support Vector Machine* Di Kota Medan” dalam penelitiannya *variabel* yang digunakan untuk proses analisis klasifikasi adalah tinggi wilayah, curah hujan, kelembapan udara, kepadatan penduduk, jumlah kasus Demam Berdarah Dengue (DBD), dan dalam Proses klasifikasi wilayah Demam berdarah memperoleh akurasi sebesar 75 %.

2 TINJAUAN PUSTAKA

2.1 Machine Learning

Machine Learning adalah metode yang digunakan untuk membuat program yang bisa belajar dari data. Berbeda dengan program komputer biasa yang statis, program machine learning adalah program yang dirancang untuk mampu belajar sendiri[5].

Cara belajar program *machine learning* mengikuti cara belajar manusia, yakni belajar dari contoh-contoh. *Machine learning* akan mempelajari pola dari contoh-contoh yang dianalisa, untuk menentukan jawaban dari pertanyaan-pertanyaan berikutnya. Memang tidak semua masalah bisa dipecahkan dengan program *machine learning*. Namun, seringkali algoritma yang sifatnya kompleks, ternyata bisa dipecahkan dengan sangat simpel oleh *machine learning*. Beberapa contoh program *machine learning* yang telah digunakan dalam kehidupan sehari-hari: Pendeteksi Spam, Pendeteksi Wajah, Rekomendasi Produk, Asisten Virtual, Diagnosa Medis, Pendeteksi Penipuan Kartu Kredit, Pengenal Digit, Perdagangan Saham dan Segmentasi Pelanggan.

2.2 Support Vector Machine

Support Vector Machine (SVM) memiliki kemampuan untuk mengatasi masalah yang bersifat linear dan nonlinear. Pendekatan kernel yang dikembangkan pada SVM juga dapat digunakan untuk mengatasi jumlah kelas yang beragam dan banyak, SVM untuk prediksi disebut dengan SVM *Regression* terdiri dari fungsi linear dan nonlinear. Metode SVM *Regression* disebut sebagai suatu teknik *nonparametric* karena bergantung pada fungsi kernel [6]. Dalam SVM ada 3 Kernel yaitu Linear, Gaussian dan Polynomial [7].

Konsep SVM dapat dijabarkan sebagai suatu cara atau metode untuk mencari *hyperplane* terbaik yang berfungsi sebagai pemisah antar dua buah *class* pada input *space* konsep tersebut memiliki prinsip *Structural Risk Minimization* atau yang biasa disebut dengan SRM, dengan tujuan menemukan *hyperplane* terbaik yang dapat memisahkan data ke dalam dua buah kelas.

2.3 k-fold Cross validation

k-fold crossvalidation merupakan suatu metode yang biasa digunakan untuk melakukan evaluasi kinerja *classifier*, metode ini dapat digunakan apabila memiliki jumlah data yang sedikit. *k-fold crossvalidation* juga biasanya digunakan untuk mengetahui rata-rata keberhasilan dari suatu sistem dengan cara melakukan subuah pengulangan dengan cara mengacak atribut masukan sehingga sistem tersebut teruji untuk beberapa atribut input yang acak yang baru [8].

2.4 Uji Koefisien Determinasi

Untuk mengetahui ketepatan atau kecocokan garis regresi yang terbentuk dalam mewakili kelompok data hasil observasi, perlu dilihat sampai seberapa jauh model yang terbentuk mampu menerangkan kondisi yang sebenarnya yang dikenal dengan nama koefisien determinasi (R^2). Nilai koefisien determinasi merupakan suatu ukuran yang menunjukkan besar sumbangan dari variabel penjelasan terhadap respon [9].

2.5 Sum Of Squares Error (SSE)

Sum Of Squares Error (SSE) adalah mengukur kesalahan penggunaan estimasi persamaan regresi untuk menghitung nilai variabel terikat dari sampel. Fungsi dari SSE dapat dilihat pada persamaan 1.

$$SSE = \sum (Y_i - \hat{Y}_i)^2 \quad (1)$$

Dengan Y_i merupakan data aktual dan \hat{Y}_i merupakan data hasil prediksi.

2.6 Total Sum Of Squares (TSS)

Total sum of Squares (TSS) adalah menggunakan nilai rata-rata (mean) hasil penjualan kuartalan dari sampel. Fungsi dari TSS dapat dilihat pada persamaan 2.

$$SST = \sum (Y_i - \bar{Y})^2 \quad (2)$$

Dengan Y_i merupakan data aktual dan \bar{Y} merupakan data *testing*.

2.7 R-squared (R^2)

R-squared (R^2) adalah suatu indikator yang digunakan untuk menggambarkan berapa banyak variasi yang dijelaskan dalam model. Berdasarkan nilai R^2 dapat diketahui tingkat signifikansi atau kesesuaian hubungan antara *variabel* bebas dan *variabel* tak bebas dalam regresi. R^2 dengan nilai mendekati 1 berarti model regresi yang telah kita buat cukup baik dalam memprediksi data dan jika nilai R^2 mendekati 0 maka model yang kita buat kurang baik dalam memprediksi data. Fungsi dari R-squared (R^2) dapat dilihat pada persamaan 3.

$$R^2 = 1 - SSE/SST \quad (3)$$

3. METODE PENELITIAN

3.1 Data

Data yang digunakan pada penelitian ini adalah data penyakit demam berdarah di kota Bandar Lampung dari tahun 2010 sampai tahun 2018, data cuaca dan juga data matriks jarak antar puskesmas yang satu dengan yang lainnya, jumlah seluruh data adalah 1.080 data dan juga memiliki 46 *variable*, yaitu : *latitude*, *longitude*, temperatur minimal, temperatur maksimal, temperatur rata-rata, curah hujan, hujan maksimal, hujan minimal, kelembapan udara rata-rata, udara maksimal, udara minimal, kecepatan angin, angin maksimal, angin minimal, jumlah penduduk, Luas Area dan juga data berupa Matriks jarak antar daerah Bakung, Beringin Raya, Campang Raya, Gedung Air, Kampung sawah, Kebon Jahe, Kedaton, Kemiling, Korpri, Kota Karang, Kupang Kota, Labuhan Ratu, Palapa, Panjang, Pasar Ambon, Permata Sukarame, Pinang Jaya, Rajabasa Indah, Satelit, Segala Mider, Simpur, Sukabumi, Sukamaju, Sukaraja, Sukarame, Sumur Batu, Susunan Baru, Way Halim, Way Kandis dan Way Laga.

3.2 Feature Dalam Data

1. *Latitude* dan *Longitude* adalah suatu sistem koordinat geografis yang digunakan untuk menentukan lokasi suatu tempat di permukaan bumi, data ini diperoleh dari *google maps*.
2. Curah hujan adalah jumlah air yang jatuh di permukaan tanah datar selama periode tertentu yang diukur dengan satuan tinggi (mm). Dalam penelitian ini data curah hujan dibagi menjadi 3, yaitu : curah hujan maksimal, minimal dan rata-rata. Data ini diperoleh dari *website* Badan Pusat Statistika(BPS) dialamat <https://lampung.bps.go.id/dynamictable/2016/11/23/71/jumlah-curah-hujan-provinsi-lampung-2010-2018.html>.
3. *Temperatur* adalah ukuran panas-dinginnya dari suatu benda. Panas-dinginnya suatu benda berkaitan dengan energi termis yang terkandung dalam benda tersebut. Makin besar energi termisnya, makin besar temperaturnya, dengan satuan (C°). Dalam penelitian ini data temperatur dibagi menjadi 3, yaitu : Temperatur maksimal, minimal dan rata-rata. Data ini diperoleh dari

website BPS dialamat <https://www.bps.go.id/statictable/2014/04/28/1347/suhuminimum-rata-rata-dan-maksimum-di-stasiun-pengamatan-bmkg-oc-2000-2018.html>.

4. *Kelembapan udara* adalah jumlah kandungan uap air yang ada dalam udara. Kandungan uap air di udara berubah-ubah bergantung suhu udara, dengan satuan (%). Dalam penelitian ini data kelembapan udara dibagi menjadi 3, yaitu : kelembapan udara maksimal, minimal dan rata-rata. Data ini diperoleh dari *website* Badan Pusat Statistika (BPS) dengan alamat <https://lampung.bps.go.id/dynamictable/2017/03/27/108/ratarata-kelembaban-udara-di-provinsi-lampung-2010-2018.html>.
5. Kecepatan angin, atau *velositas* gelombang angin, adalah sebuah kuantitas atmosferik fundamental. Kecepatan angin disebabkan oleh pergerakan angin dari tekanan tinggi ke tekanan rendah, dengan satuan (Knot). Dalam penelitian ini data Kecepatan angin dibagi menjadi 3, yaitu : Kecepatan angin maksimal, minimal dan rata-rata. Data ini diperoleh dari *website* Badan Pusat Statistika (BPS) <https://lampung.bps.go.id/dynamictable/2017/03/27/107/ratarata-kecepatan-angin-di-provinsi-lampung-2010-2018.html>.
6. *Jumlah penduduk* adalah jumlah manusia yang bertempat tinggal atau berdomisili pada suatu wilayah atau daerah dan memiliki mata pencaharian tetap. Data ini diperoleh dari BPS dialamat <https://bandarlampungkota.bps.go.id/statictable/2015/12/08/19/jumlah-penduduk-kota-bandar-lampung-2010-2020.html>.
7. Matriks jarak adalah jarak lokasi antar puskesmas yang satu dengan puskesmas yang lainnya, dengan satuan (Km²) Data ini diperoleh dari perhitungan *Latitude* dan *Longitude*.

3.3 Alat

Dalam penelitian ini peneliti menggunakan *Software* dan *Hardware* sebagai berikut :

1. Perangkat Lunak

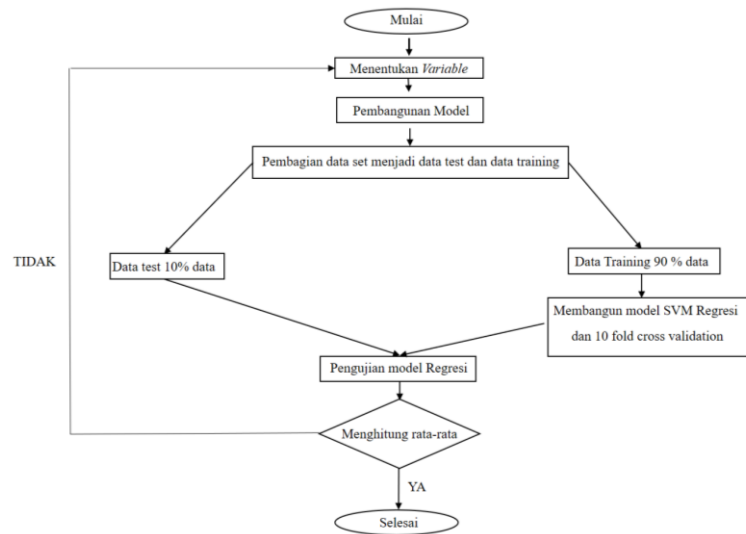
Sistem operasi Windows 10 Profesional 64 bit, Microsoft Excel 2010, R studio versi 3.5.1 64 bit, *Package caret* versi 6.0-81, *Package kernlab* versi 0.9-27 dan R versi 3.5.1 64 bit.

2. Perangkat Keras

Prosesor Intel® Core™ i5-7200U (2.0 GHz),4 Threads, 2 Cores, Cache, 4 GT/s OPI Bus Speed,2.5-3.1GHz CPU Clock Speed, VGA NVIDIA® Geforce® 940Mx, DDR4-2133, RAM 8GB, Memory 128 GB SSD + 1TB HDD, Max Memory Bandwidth 34.1 GB/s.

3.4 Tahap Penelitian

Prediksi persebaran demam berdarah di Kota Bandar Lampung ini berlangsung dalam beberapa tahap, yaitu menentukan *variable* dengan cara memilih *variable* apa saja yang sesuai dengan prediksi demam berdarah, tahapan selanjutnya ialah membangun model SVM regresi, setelah membangun model svm dilanjutkan dengan pembagian data menjadi dua bagian, yaitu data *testing* dan data *training* menggunakan 10-fold cross validation sebanyak 10 kali dengan aturan 10% untuk data *testing* dan 90% untuk data *training*, setelah pembagian data dilanjutkan dengan proses prediksi data *training* dengan data *testing* yang telah dibuat sebelumnya, setelah dilakukan prediksi data maka proses terakhir ialah perhitungan R² dari setiap percobaan. Tahapan penelitian tersebut ditunjukkan pada Gambar 1.



Gambar 1. Metode Penelitian Prediksi Demam Berdarah

3.5 Pengumpulan Data

Penelitian ini menggunakan Data Demam Berdarah di Kota Bandar Lampung pada tahun 2010 sampai tahun 2018, selain menggunakan data penderita DBD penelitian ini juga menggunakan data cuaca yang di peroleh dari website Badan Pusat Statistika(BPS) dan matrix jarak yang diperoleh dari perhitungan jarak antara lokasi penderita 1 dengan penderita yang lain berdasarkan letak *Latitude* dan *Longitude* nya.

3.6 Pembangunan Model

Penelitian ini dilakukan dengan model *Support Vector Machine* (SVM) dengan menggunakan 3 kernel, yaitu Linear, Gaussian dan polynomial [11]. Pada persamaan 4 meunjukkan fungsi kernel Linier, persamaan 5 menunjukkan fungsi kernel Gaussian dan persamaan 6 menunjukkan fungsi kernel Polynomial.

$$K(x_i, x_j) = x_i^T x_j + C \quad (4)$$

Dengan X_i dan X_j merupakan vektor dari data set dan C merupakan constant.

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (5)$$

Dengan X_i dan X_j merupakan vektor dari data set, γ adalah parameter untuk mengkontrol kecepatan proses learning dan \exp merupakan basis dari logaritma alami.

(6)

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$$

Dengan X_i dan X_j merupakan vektor dari data set, γ adalah parameter untuk mengkontrol kecepatan proses learning dan d merupakan pangkat polynomial yang digunakan.

3.7 Pengujian Model

Tahap ini dilakukan ketika sudah mendapatkan model SVM terbaik. Model SVM tersebut akan digunakan untuk melakukan pengujian dengan data testing yang telah kita miliki dari proses pembagian dataset, *k-fold cross validation* merupakan suatu teknik yang dapat digunakan apabila memiliki jumlah data yang terbatas (jumlah instance tidak banyak) *k-fold cross validation* merupakan salah satu metode

Percobaan Dengan Matrix Jarak Untuk Kernel Linear, Gaussian dan Polynomial

Pada percobaan kedua yang peneliti lakukan ialah percobaan dengan menggunakan SVM dengan kernel Linear, Gaussian dan Polynomial. Dalam percobaan ini jumlah data yang digunakan adalah 1.080 dengan 44 Variable. Dari percobaan yang dilakukan pada kernel Gaussian mendapatkan R^2 tertinggi yaitu sebesar 75,52%, sedangkan R^2 pada kernel Linear mendapatkan R^2 terendah dari kernel yang lain yaitu sebesar 74,61% dan pada Kernel Polynomial mendapatkan R^2 sebesar 75,15%, seperti yang dapat dilihat pada Table 4.

Tabel 1. Hasil Percobaan Dengan Metode SVM Linear, Gaussian dan Polynomial Tanpa Matriks Jarak.

Type	10-fold cv	Evaluasi pengujian		
		SSE	TSS	Rsquared
SVM Linear	Rata-rata	1550,1332	5099,4177	70,00
SVM Gaussian	Rata-rata	1287,2395	4859,2295	70,79
SVM Polynomial	Rata-rata	1583,6607	5400,3322	70,61

Tabel 2. Hasil Percobaan Dengan Metode SVM Linear, Gaussian dan Polynomial Dengan Matriks Jarak.

Type	10-fold cv	Evaluasi pengujian		
		SSE	TSS	Rsquared
SVM Linear	Rata-rata	1196,5	4649,454	74,41
SVM Gaussian	Rata-rata	1169,4	4442,0244	75,44
SVM Polynomial	Rata-rata	1189,4	4779,45	75,11

Tabel 3. Hasil Percobaan Dengan Metode SVM Linear, Gaussian dan Polynomial Tanpa Matriks Jarak.

Type	10-fold cv	Evaluasi pengujian		
		SSE	TSS	Rsquared
SVM Linear	Rata-rata	4103,6007	4523,7582	74,13
SVM Gaussian	Rata-rata	2033,4893	4510,1349	74,63
SVM Polynomial	Rata-rata	1311,2787	5186,1797	74,59

Tabel 4. Hasil Percobaan Dengan Metode SVM Linear, Gaussian dan Polynomial Tanpa Matriks Jarak.

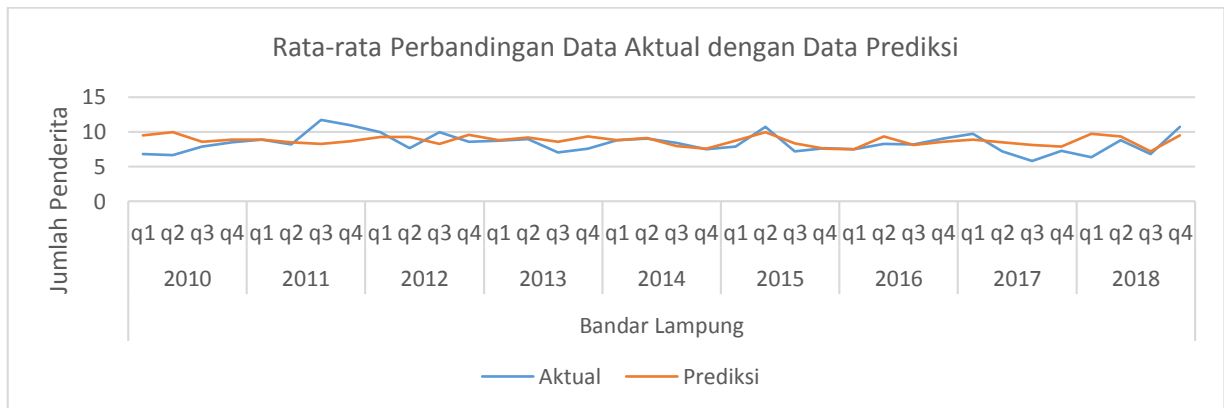
Type	Iterasi	Evaluasi pengujian		
		SSE	TSS	Rsquared
SVM Linear	Rata-rata	2725,3704	4640,6035	74,61
SVM Gaussian	Rata-rata	2363,3177	5730,061325	75,52
SVM Polynomial	Rata-rata	2311,0466	4483,9521	75,15

4 PEMBAHASAN

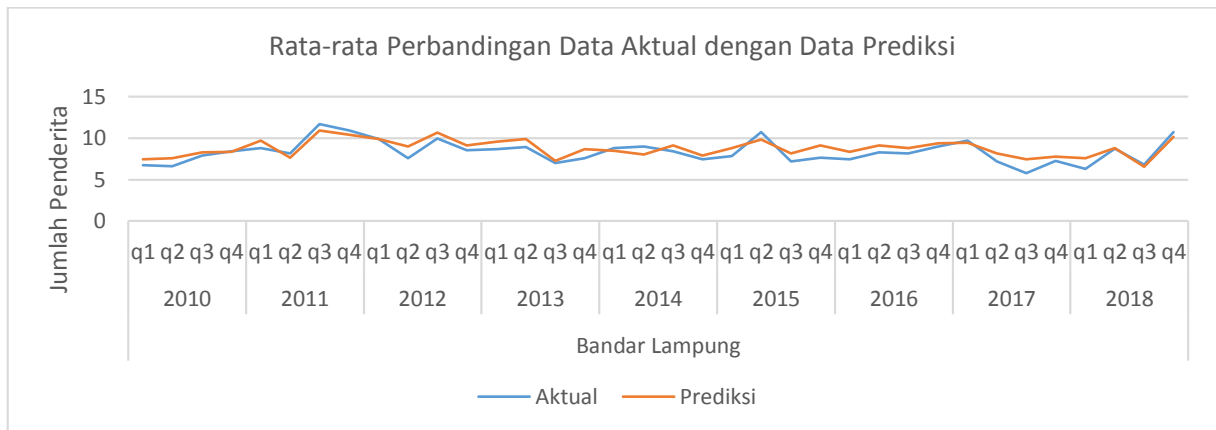
Dalam Penelitian ini dilakukan empat kali percobaan, dua percobaan pertama dilakukan tanpa melakukan *feature selection*, dua percobaan tersebut ialah percobaan tanpa matriks jarak dan percobaan dengan matriks jarak, dalam percobaan dengan menggunakan matriks jarak data yang digunakan sebanyak 1.080 dengan 10 *variabel*. Dengan data tersebut didapatkan R^2 sebesar 70,00% untuk *kernel* Linear, 70,79% untuk *kernel* Gaussian dan 70,61% untuk *kernel* Polynomial. Percobaan yang ke dua ialah percobaan dengan menggunakan matriks jarak, dalam percobaan ini jumlah data yang digunakan sebanyak 1.080 data dengan 46 *variable*. Dengan data tersebut didapatkan R^2 sebesar 74,40% untuk *kernel* Linear, 75,44% untuk *kernel* Gaussian dan 75,10% untuk *kernel* Polynomial.

Setelah dua percobaan pertama selesai tanpa melakukan *feature selection*. dilanjutkan percobaan dengan melakukan *feature selection* dalam percobaan ini dilakukan dua kali percobaan juga yaitu percobaan tanpa Matriks jarak dan percobaan dengan menggunakan matriks jarak. Dalam percobaan tanpa menggunakan Matriks jarak jumlah data yang di gunakan sebanyak 1.080 data dengan 8 *variable*. Dengan data tersebut didapatkan R^2 sebesar 74,12% untuk *kernel* Linear, 74,62% untuk *kernel* Gaussian dan 74,58% untuk *kernel* Polynomial. Percobaan yang ke dua ialah percobaan dengan menggunakan matriks jarak, dalam percobaan ini jumlah data yang digunakan sebanyak 1.080 data dengan 44. Dengan data tersebut didapatkan R^2 sebesar 74,61% untuk *kernel* Linear, 75,51% untuk *kernel* Gaussian dan 75,15% untuk *kernel* Polynomial.

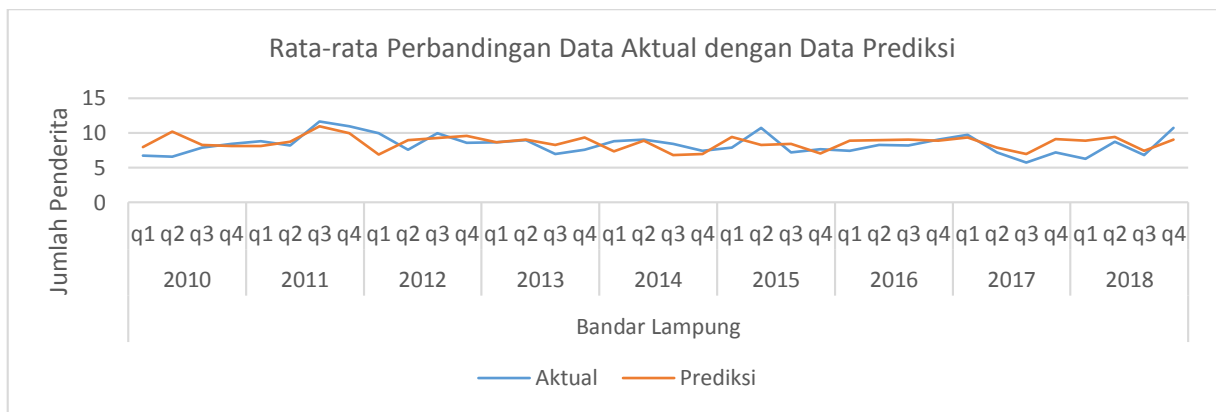
Pada Gambar 3 menunjukkan perbandingan data hasil prediksi dengan data aktual untuk percobaan dengan matriks jarak yang diterapkan *feature selection* dengan tingkat R^2 74,61%, Gambar 4 menunjukkan perbandingan data hasil prediksi dengan data aktual untuk percobaan dengan matriks jarak yang diterapkan *feature selection* dengan tingkat R^2 75,51% dan Gambar 5 menunjukkan perbandingan data hasil prediksi dengan data aktual untuk percobaan dengan matriks jarak yang diterapkan *feature selection* dengan tingkat R^2 75,15%.



Gambar 3. Perbandingan Data Aktual dengan Data Hasil Prediksi dengan Metode SVM Linear.



Gambar 4. Perbandingan Data Aktual dengan Data Hasil Prediksi dengan Metode SVM Gaussian.



Gambar 5. Perbandingan Data Aktual dengan Data Hasil Prediksi dengan Metode SVM Polynomial.

5 PERBANDINGAN

Penelitian serupa pernah dilakukan oleh Harold Situmorang dalam Tesisnya yang berjudul “Klasifikasi Wilayah Epidemik Demam Berdarah Menggunakan Metode Support Vector Machine Di Kota Medan” [13]. Data yang digunakan dalam penelitiannya adalah data jumlah kasus Demam Berdarah di kota Medan mulai tahun 2010-2013 yang diperoleh dari dinas Kesehatan Kota Medan dan dibagi dalam 21 Kecamatan. Pemilihan variabel penelitian dalam klasifikasi wilayah epidemik Demam Berdarah didasarkan pada data spasial kerentanan penyakit Demam Berdarah di Kota Medan. Dalam penelitiannya *variabel* yang digunakan untuk proses analisis klasifikasi adalah tinggi wilayah, curah hujan, kelembapan udara, kepadatan penduduk dan jumlah kasus Demam Berdarah Dengue (DBD) untuk daerah Medan, dan dalam Proses klasifikasi wilayah Demam berdarah memperoleh akurasi pengujian sebanyak 75%.

Sedangkan dalam penelitian ini, dilakukan 4 kali percobaan yaitu dengan menggunakan data dari tahun 2010 sampai tahun 2018 dengan jumlah data sebanyak 1.080 data. Dalam empat kali percobaan tersebut *variable* yang digunakan adalah sebanyak 10 *variable*, 46 *variable*, 8 *variable* dan 44 *variable*, *variable* tersebut adalah latitude, longitude, tempmin, tempmax, tempratarata, Curahhujan, hujanmax, hujanmin, kelembabanudararata, udaramax, udaramin, kecepatanangin, anginmax,

anginmin, jmlhpnddk, LuasArea , Bakung, BeringinRaya, CampangRaya, GedungAir, Kampungsawah, KebonJahe, Kedaton, Kemiling, Korpri, KotaKarang, KupangKota, LabuhanRatu, Palapa, Panjang, PasarAmbon, PermataSukarame, PinangJaya, RajabasaIndah, Satelit, SegalaMider, Simpur, Sukabumi, Sukamaju, Sukaraja, Sukarame, Sumurbatu , SusunanBaru, WayHalim, Wayandis, WayLaga. Hasil dari percobaan yang saya lakukan dapat dilihat pada Tabel 5

Tabel 5. Hasil Perbandingan Percobaan dan perbandingan dengan penelitian oleh Situmorang 2015.

	Tanpa Feature Selection		Dengan Feature Selection		Situmorang
	10 var	46 var	8 var	44 var	
Linear	70,00%	74,40%	74,12%	74,61%	
Gaussian	70,79%	75,44%	74,62%	75,51%	75,00%
Polynomial	70,61%	75,10%	74,58%	75,15%	

6 KESIMPULAN

Dari hasil penelitian yang dilakukan, dapat diambil simpulan bahwa telah dibangun sebuah model prediksi demam berdarah di Kota Bandar Lampung dengan menggunakan metode *Support Vector Machine*, dengan menggunakan 3 kernel, yaitu : Linear, Gaussian dan Polynomial. Setelah dilakukan empat kali percobaan pada Percobaan tanpa melakukan melakukan Feature Selection didapatkan tingkat R^2 sebesar 70,00% untuk kernel Linear, 70,79% untuk kernel Gaussian dan 70,61% untuk kernel Polynomial. Untuk percobaan tanpa matrix jarak, R^2 sebesar 74,40% untuk kernel Linear, 75,44% untuk kernel Gaussian dan 75,10% untuk kernel Polynomial. Untuk Percobaan dengan menggunakan matrix jarak sedangkan pada percobaan dengan melakukan *Feature Selection* didapatkan R^2 sebesar 74,12% untuk kernel Linear, 74,62% untuk kernel Gaussian dan 74,58% untuk kernel Polynomial. Untuk percobaan tanpa matrix jarak dan R^2 sebesar 74,61% untuk kernel Linear, 75,51% untuk kernel Gaussian dan 75,15% untuk kernel Polynomial. Untuk percobaan dengan menggunakan matrix jarak. Dengan penambahan data matrix jarak pada setiap percobaan dapat meningkatkan R^2 pada setiap kernel.

DAFTAR PUSTAKA

- [1] BPS.Rara-rataKelembapanUdara. <https://lampung.bps.go.id/dynamictable/2017/03/27/108/ratarata-kelembaban-udara-di-provinsi-lampung-2010-2018.html>. Diakses pada 03 Maret 2019.
- [2] BPS. Jumlah Curah Hujan. <https://lampung.bps.go.id/dynamictable/2016/11/23/71/jumlah-curah-hujan-provinsi-lampung-2010-2018.html>.Diakses pada 03 MArset 2019
- [3] BPS. Rata-rata Temperatur. <https://www.bps.go.id/statictable/2014/04/28/1347/suhuminimum-rata-rata-dan-maksimum-di-stasiun-pengamatan-bmkg-oc-2000-2018.html>. Diakses pada 03 Maret 2019.
- [4] BPS. Rata-rata Kecepatan Angin. <https://lampung.bps.go.id/dynamictable/2017/03/27/107/ratarata-kecepatan-angin-di-provinsi-lampung-2010-2018.html>. Diakses pada 03 Maret 2019.BPS. Jumlah Penduduk Bandar

Lampung. <https://bandarlampungkota.bps.go.id/statictable/2015/12/08/19/jumlah-penduduk-kota-bandar-lampung-2010-2020.html>. Diakses pada 03 Maret 2019.

- [5] Fanany, D. (2012). Dengue hemorrhagic fever and natural disaster: the case of Padang, West Sumatra. *International Journal of Collaborative Research on Internal Medicine and Public Health (IJCRIMPH)*, 4(5), 673–678.
- [6] Fauzi, Nungsiyati, Noviarti, T., Muslihudin, M., Irviani, R., & Maseleno, A. (2018). Optimal Dengue Endemic Region Prediction using Fuzzy Simple Additive Weighting based Algorithm. *International Journal of Pure and Applied Mathematics*, 118(7 Special Issue).
- [7] Setyawan, E., Subantoro, R., & Prabowo, R. (2016). Analisis Peramalan (Forecasting) Produksi Karet (*Hevea Brasiliensis*) Di PT Perkebunan Nusantara IX Kebun Sukamangli Kabupaten Kendal. *Analisis Peramalan (Forecasting)*, 12(2), 11–19.
- [8] Kusumodestoni, R. H., & Sarwido, S. (2017). Komparasi Model Support Vector Machines (Svm) Dan Neural Network Untuk Mengetahui Tingkat Akurasi Prediksi Tertinggi Harga Saham. *Jurnal Informatika Upgris*, 3(1).
- [9] W. Yuciana. (2014). Penerapan Metode Klasifikasi Support Vector Machine (Svm) Pada Data Akreditasi Sekolah Dasar (SD) Di Kabupaten Magelang, 3(8), 811–820.
- [10] Nurmasani, A. (2017). Analisis Support Vector Machine. *Jurnal Informasi Interaktif*, 2(1).
- [11] Bhavsar, H., & Panchal, M. H. (2012). A Review on Support Vector Machine for Data Classification, 1(10), 185–189.
- [12] Banjarsari, M. A., Budiman, I., & Farmadi, A. (2016). Penerapan K-Optimal Pada Algoritma KNN Untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan IP Sampai Dengan Semester 4.
- [13] Situmorang, H. 2015. Klasifikasi Wilayah Demam Berdarah dengan Metode SVM di Kota Medan. Tesis Magister. Universitas Sumatra Utara