

PREDIKSI JUMLAH PENDERITA PENYAKIT TUBERKULOSIS DI KOTA BANDAR LAMPUNG MENGUNAKAN METODE SVM (*Support Vector Machine*)

¹Favorisen R. Lumbanraja, ²Ira Hariati Br Sitepu, ³Didik Kurniawan,
⁴Aristoteles.

Jurusan Ilmu Komputer, FMIPA,
Universitas Lampung, Bandar Lampung, 35141

¹favorisen.lumbanraja@fmipa.unila.ac.id, ²ira.hariatii@gmail.com,
³didik.kurniawan@fmipa.unila.ac.id, ⁴aristoteles.1981@fmipa.unila.ac.id

Abstrak

Tuberkulosis (TB atau TBC) merupakan salah satu penyakit infeksi yang disebabkan oleh Bakteri *Mycobacterium tuberculosis*. Bakteri tersebut merupakan bakteri yang sangat kuat sehingga dalam pengobatannya memerlukan waktu yang cukup lama. Pengobatan penyakit tuberkulosis dilakukan selama 6-9 bulan secara rutin dengan sedikitnya 3 macam jenis obat. Saat ini kebanyakan masyarakat menganggap batuk dalam jangka waktu berbulan-bulan merupakan batuk biasa, jika dicermati salah satu gejala yang ditimbulkan penyakit tuberkulosis, yaitu batuk dalam jangka waktu yang panjang. Pada penelitian ini digunakan data penderita tuberkulosis di Kota Bandar Lampung, data cuaca dan matrix jarak antara kejadian penderita tuberkulosis yang satu dengan kejadian yang lainnya dalam lingkup kecamatan. Jumlah dari keseluruhan data sebanyak 600 data dengan 44 variabel. Penelitian ini juga menggunakan 3 kernel yaitu, Linear, Gaussian, dan Polynomial dengan menggunakan Metode SVM dengan kernel Linear mendapatkan nilai rata-rata R^2 sebesar 51.43 %, pada percobaan dengan metode SVM dengan kernel Gaussian mendapatkan nilai rata-rata R^2 sebesar 58.53 % dan pada percobaan dengan metode SVM dengan kernel Polynomial mendapatkan nilai rata-rata R^2 sebesar 36.03 %.

Kata Kunci : Prediksi penderita tuberkulosis, tuberkulosis, Machine Learning, Support Vector Machine.

Abstract

Tuberculosis (TB / TBC) is one of infectious disease caused by *Mycobacterium tuberculosis* bacteria. These bacteria are very strong bacteria so for the treatment takes a long time. Tuberculosis treatment is carried out for 6-9 months regularly with at least 3 types of drugs. Currently, most of people consider a cough for months is a common cough, if looked by one of the symptoms caused by tuberculosis, which is a cough for a long time. In this research, data on tuberculosis patients in the city of Bandar Lampung were used, weather data and the distance matrix between the case of tuberculosis patients with other case within the district. The total number of data is 600 data with 44 variables. This research also uses 3 kernels namely, Linear, Gaussian, and Polynomial by using the SVM method with the Linear kernel getting an average R^2 value of 51.43%, in the experiment with the SVM method with a gaussian

kernel getting an average R^2 value of 58.53% and at Experiments with the SVM method with the Polynomial kernel obtained an average value of R^2 of 36.03% .

Keywords : *Prediction of tuberculosis sufferers, tuberculosis, Machine Learning, Support Vector Machine.*

1. PENDAHULUAN

Tuberculosis adalah salah satu penyakit infeksi yang disebabkan oleh Bakteri *Mycobacterium tuberculosis*. Menurut WHO *Global Surveillance* penyakit tuberkulosis di dunia pada tahun 2014 mencapai 5.4 juta penderita dan membunuh 1.5 juta penderita sehingga penyakit ini menjadi salah satu penyakit terbesar di dunia [1]. Dalam tubuh manusia bakteri ini lebih sering menimbulkan infeksi pada organ paru-paru dibandingkan dengan bagian organ tubuh lainnya yang dikenal dengan TB Paru (Tuberkulosis Paru). Kasus penyakit tuberkulosis paru sering terjadi di Indonesia, bahkan diseluruh dunia penyakit TB paru ini menjadi masalah utama kesehatan masyarakat. Penyakit ini dapat menular melalui udara saat penderita bersin atau batuk [2] Di Indonesia pada tahun 2015 penderita tuberkulosis mencapai 302.411 orang (Laki-laki 177.914 orang dan Perempuan 124.497 orang) dan menempati peringkat ke-4 di dunia dengan penderita tuberkulosis terbanyak [3].

Metode *Support Vector Machine* (SVM) banyak digunakan dalam penelitian-penelitian untuk memprediksi suatu kasus tertentu. *Support Vector Machine* (SVM) dikenalkan pertama kali oleh Vapnik tahun 1992 sebagai salah satu metode *learning machine* yang bekerja dengan prinsip *Structural Risk Minimization* (SRM) yang bertujuan untuk menemukan *hyperplane* untuk memisahkan dua buah *class* pada *input space*. Metode ini menggunakan hipotesis berupa fungsi-fungsi *linear* dalam sebuah ruang fitur yang berdimensi tinggi, dengan mengimplementasikan *learning bias* yang berasal dari teori pembelajaran statistik. Tingkat akurasi pada model yang akan dihasilkan oleh proses peralihan dengan SVM sangat bergantung terhadap fungsi kernel dan parameter yang digunakan [4]. SVM memiliki banyak fitur penting yang didukung untuk menghasilkan hasil empiris yang baik dan mendukung untuk klasifikasi dan estimasi fungsi non-linear [5].

Penelitian sebelumnya pernah dilakukan oleh Lumbanraja tahun 2019 untuk memprediksi jumlah penderita demam berdarah di kota Bandar Lampung dengan Metode *Support Vector Machine* [6]. Dalam penelitiannya penulis melakukan empat kali percobaan dengan 10 variabel, 46 variabel, 8 variabel, dan 44 variabel, variable tersebut adalah latitude, longitude, tempmin, tempmax, temperatur rata-rata, Curahhujan, hujanmax, hujanmin, kelembabanudara rata, udaramax, udaramin, kecepatan angin, anginmax, anginmin, jumlah penduduk, Luas Area, 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, Simpung, Sukabumi, Sukamaju, Sukaraja, Sukarame, Sumurbatu, Susunan Baru, Way Halim, Way Kandis, Way Laga. Dengan memperoleh akurasi pengujian yaitu

Percobaan dengan menggunakan Metode SVM dengan kernel Linear mendapatkan nilai rata-rata R^2 sebesar 74,41%, pada percobaan metode dengan SVM kernel gaussian mendapatkan nilai rata-rata R^2 75,44% dan pada percobaan SVM Polynomial mendapatkan nilai rata-rata R^2 sebesar 75,11% .

Pada penelitian kali ini berupaya untuk membuat model prediksi dengan menggunakan metode *Support Vector Machine* dengan 3 kernel yaitu, Linear, Gaussian, dan Polynomial.

2. TINJAUAN PUSTAKA

2.1. *Machine Learning*

Machine Learning adalah pembelajaran mengenai algoritma dalam mempelajari sesuatu untuk melakukan beberapa hal tertentu yang secara otomatis dilakukan manusia. *Machine Learning* adalah salah satu bidang kecerdasan buatan yang bisa mempengaruhi berbagai aspek lainnya, yaitu matematika, statistika, dan berbagai aspek dari komputer sains. *Machine Learning* bertujuan untuk mempelajari sebuah algoritma dalam melakukan sistem belajar secara otomatis dengan kontribusi yang sangat minimal yang pada umumnya dilakukan oleh manusia [xxxx].

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. Tidak semua masalah bisa dipecahkan dengan program *Machine Learning*. Namun, sering kali algoritma yang sifatnya kompleks, ternyata bisa diselesaikan dengan sangat simpel oleh *Machine Learning*.

2.2. *Support Vector Machine*

SVM (*Support Vector Machine*) dikenalkan pertama kali oleh Vapnik tahun 1992 sebagai salah satu metode learning machine yang bekerja dengan prinsip *Structural Risk Minimization* (SRM) yang bertujuan untuk menemukan hyperplane untuk memisahkan dua buah class pada input space. Metode ini menggunakan hipotesis berupa fungsi – fungsi linier dalam sebuah ruang fitur yang berdimensi tinggi, dengan mengimplementasikan learning bias yang berasal dari teori pembelajaran statistik. Tingkat akurasi pada model yang akan dihasilkan oleh proses peralihan dengan SVM sangat bergantung terhadap fungsi kernel dan parameter yang digunakan [xxx]. SVM merupakan sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi-fungsi linier dalam sebuah ruang fitur berdimensi tinggi, dilatih dengan algoritma pembelajaran yang didasarkan pada teori optimasi dengan mengimplementasikan learning bias yang berasal dari teori pembelajaran statistik [7]

2.3. *Evaluasi Regresi*

Untuk mengetahui ketepatan untuk garis regresi yang terbentuk untuk mewakili kelompok data hasil observasi, sampai seberapa jauh model yang terbentuk mampu menerangkan kondisi yang sebenarnya dikenal dengan nama koefisien determinasi (R^2) [8].

a. *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.

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

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

b. *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.

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

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

c. *R-Squared (R²)*

R-squared (R²) adalah fungsi untuk mengukur tingkat keberhasilan model regresi yang kita gunakan dalam memprediksi nilai. R^2 dengan nilai mendekati 1 berarti data yang cukup bermacam-macam sudah di tangani dengan baik oleh model regresi dan jika 0 maka sebaliknya. Fungsi dari *R-Squared* dapat dilihat pada persamaan 3.

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

3. METODE PENELITIAN

3.1 Data

Penelitian ini menggunakan Data Tuberculosis di Kota Bandar Lampung pada tahun 2014 sampai tahun 2018, selain data penderita tuberculosis pada penelitian ini juga menggunakan data cuaca yang di dapatkan dari website Badan Pusat Statistika (BPS) dan matrix jarak yang diperoleh dari perhitungan jarak antara lokasi penderita yang 1 dengan penderita yang lainnya berdasarkan kecamatan dan letak *latirude* dan *longitude* nya.

3.2 Pembangunan Model

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

$$K(X_i, X_j) = X_i \cdot X_j \dots\dots\dots(4)$$

Dengan X_i dan X_j merupakan vektor dari data set [9]

$$K(X_i, X_j) = e^{-\frac{\|X_i - X_j\|^2}{2\sigma^2}} \dots\dots\dots(5)$$

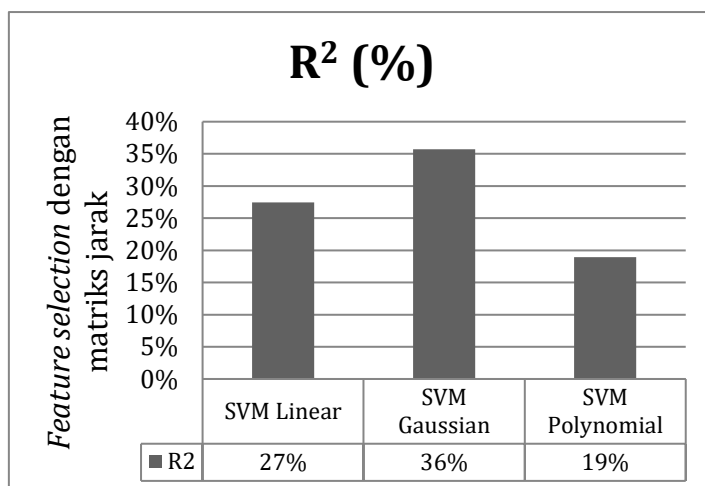
Ratu, Palapa, Panjang, Pasar Ambon, Permata Sukarame, Pinang Jaya, Rajabasa Indah, Satelit, Segala Mider, Simpung, Sukabumi, Sukamaju, Sukaraja, Sukarame, Sumur batu, Susunan baru, Way Halim, Way Kandis dan Way Laga). Proses pembagian data set dilakukan dengan aturan 90% data dijadikan sebagai *training data* dan 10% data dijadikan sebagai *test data*.

4.1.1 Percobaan dengan *Feature Selection* dengan Matriks

Pada penelitian dengan menggunakan *feature selection* didapatkanlah data sebanyak 600 data dengan atribut sebanyak 42 atribut, yang terdiri dari : penderita, latitude, longitude, curah hujan, hujan maksimum, hujan minimum, kelembaban udara rata-rata, udara maksimum, udara minimum, kecepatan angin, angin maksimum, angin minimum, jarak antar ke- 30 puskesmas (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, Simpung, Sukabumi, Sukamaju, Sukaraja, Sukarame, Sumur batu, Susunan baru, Way Halim, Way Kandis dan Way Laga). Selain menggunakan data tersebut penelitian ini juga menggunakan 3 kernel yaitu, Linear, Gaussian, dan Polynomial. Pada Table 1. menunjukkan hasil percobaan dengan menggunakan kernel Linier, kernel Gaussian dan kernel Polynomial.

Tabel 1. Hasil Percobaan dengan *Feature Selection* dengan Matriks Jarak Untuk *Kernel Linear, Gaussian, dan Polynomial* dengan Metode SVM.

Type	Iterasi	Evaluasi pengujian		
		SSE	TSS	R ²
SVM Linear	Rata-rata (dari 10 percobaan)	3902.1526	4728.1207	0.27471274
SVM Gaussian	Rata-rata (dari 10 percobaan)	3697.9327	5652.329	0.3570446
SVM Polynomial	Rata-rata (dari 10 percobaan)	5693.8487	7246.2122	0.18937049



Gambar 2. Hasil Percobaan dengan *Feature Selection* dengan Matriks Jarak.

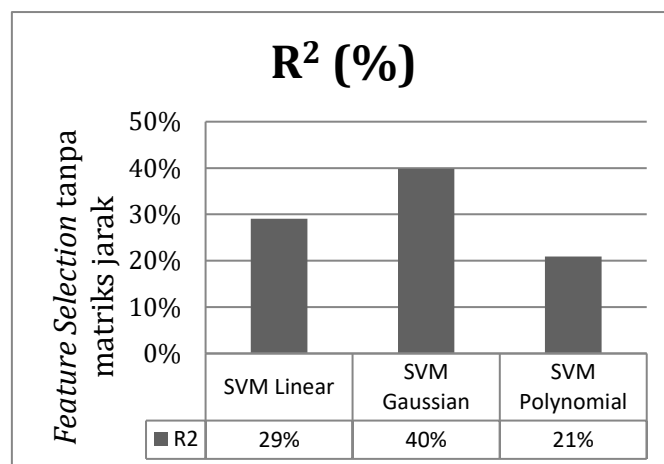
Dari percobaan yang dilakukan diatas didapatkanlah R² tertinggi pada *kernel Gausiiian atau Radial* dengan R² sebesar 35.70 %, sedangkan pada *kernel Linear* menghasilkan R² sebesar 27.47 %, dan R² pada *kernel Polynomial* merupakan R² terendah pada percobaan ini dengan nilai R² sebesar 18.94 %.

4.1.2 Percobaan dengan *Feature Selection* tanpa Matriks

Pada penelitian dengan menggunakan *feature selection* didapatkanlah data sebanyak 600 data dengan atribut sebanyak 12 atribut, yang terdiri dari : penderita, latitude, longitude, curah hujan, hujan maksimum, hujan minimum, kelembaban udara rata-rata, udara maksimum, udara minimum, kecepatan angin, angin maksimum, angin minimum. Selain menggunakan data tersebut penelitian ini juga menggunakan 3 kernel yaitu, Linear, Gaussian, dan Polynomial. . Pada Table 2 menunjukkan hasil percobaan dengan menggunakan kernel Linier, kernel Gaussian dan kernel Polynomial.

Tabel 2. Hasil Percobaan dengan *Feature Selection* tanpa Matriks Jarak Untuk *Kernel Linear, Gaussian, dan Polynomial* dengan Metode SVM.

Type	Iterasi	Evaluasi pengujian		
		SSE	TSS	R ²
SVM Linear	Rata-rata (dari 10 percobaan)	4553.1315	5432.658	0.29088119
SVM Gaussian	Rata-rata (dari 10 percobaan)	3317.4852	5420.445	0.39827412
SVM Polynomial	Rata-rata (dari 10 percobaan)	5257.7827	6649.7193	0.20932436



Gambar 3. Hasil Percobaan dengan *Feature Selection* tanpa Matriks Jarak.

Dari percobaan yang dilakukan diatas didapatkanlah R² tertinggi pada *kernel Gausiiian atau Radial* dengan R² sebesar 39.83 %, sedangkan pada *kerner Linear* menghasilkan akurasi sebesar 29.09 %, dan R² pada *kernel Polynomial* merupakan R² terendah pada percobaan ini dengan nilai R² sebesar 20.93 %.

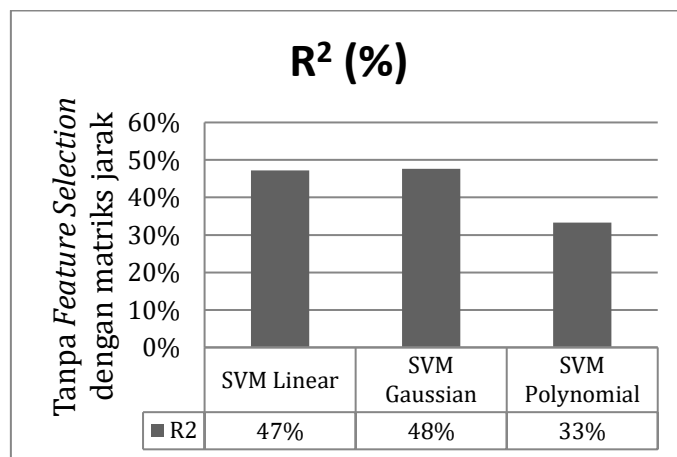
4.1.3 Percobaan dengan Tanpa *Feature Selection* dengan Matriks

Pada penelitian dengan menggunakan *feature selection* didapatkanlah data sebanyak 600 data dengan atribut sebanyak 44 atribut, yang terdiri dari : penderita, latitude, longitude, curah hujan, hujan maksimum, hujan minimum, kelembaban udara rata-rata, udara maksimum, udara minimum, kecepatan angin, angin maksimum, angin minimum, jumlah penduduk, luas area jarak antar ke- 30 puskesmas (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). Selain menggunakan data tersebut penelitian ini juga menggunakan 3 kernel yaitu, Linear, Gaussian, dan Polynomial. Pada Table 3 menunjukan hasil percobaan dengan menggunakan kernel Linier, kernel Gaussian dan kernel Polynomial.

Tabel 3. Hasil Percobaan tanpa *Feature Selection* dengan Matriks Jarak Untuk *Kernel Linear, Gaussian, dan Polynomial* dengan Metode SVM.

Type	Iterasi	Evaluasi pengujian		
		SSE	TSS	R ²
SVM Linear	Rata-rata (dari 10 percobaan)	3373.9551	6609.9256	0.47219411
SVM Gaussian	Rata-rata (dari 10 percobaan)	3580.3131	6781.2416	0.4766504

SVM Polynomial	Rata-rata (dari 10 percobaan)	4949.5601	7233.0301	0.3328422
----------------	-------------------------------	-----------	-----------	------------------



Gambar 4. Hasil Percobaan tanpa *Feature Selection* dengan Matriks Jarak.

Dari percobaan yang dilakukan diatas didapatkanlah R² tertinggi pada *kernel Gausiiian atau Radial* dengan R² sebesar 47.67 % , sedangkan pada *kerner Linear* menghasilkan R² sebesar 47.22 % , dan akurasi pada *kernel Polynomial* merupakan akurasi terendah pada percobaan ini dengan nilai akurasi R² 33.28%.

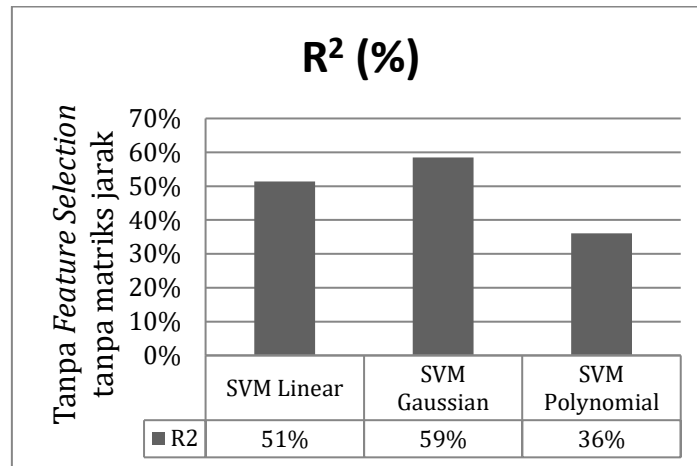
4.1.4 Percobaan dengan Tanpa *Feature Selection* tanpa Matriks

Pada penelitian dengan menggunakan *feature selection* didapatkanlah data sebanyak 600 data dengan atribut sebanyak 14 atribut, yang terdiri dari: penderita, latitude, longitude, curah hujan, hujan maksimum, hujan minimum, kelembaban udara rata-rata, udara maksimum, udara minimum, kecepatan angin, angin maksimum, angin minimum, jumlah penduduk, dan luas area. Selain menggunakan data tersebut penelitian ini juga menggunakan 3 kernel yaitu, Linear, Gaussian, dan Polynomial. . Pada Table 4 menunjukan hasil percobaan dengan menggunakan kernel Linier, kernel Gaussian dan kernel Polynomial.

Tabel 4. Hasil Percobaan tanpa *Feature Selection* tanpa Matriks Jarak Untuk *Kernel Linear, Gaussian, dan Polynomial* dengan Metode SVM.

Type	Iterasi	Evaluasi pengujian		
		SSE	TSS	R ²
SVM Linear	Rata-rata (dari 10 percobaan)	2966.854	6058.8518	0.51431736
SVM Gaussian	Rata-rata (dari 10 percobaan)	2853.6755	6866.4143	0.58525575

SVM Polynomial	Rata-rata (dari 10 percobaan)	4107.5629	6339.1651	0.36033585
-------------------	-------------------------------------	-----------	-----------	-------------------



Gambar 5. Hasil Percobaan tanpa *Feature Selection* tanpa Matriks Jarak.

Dari percobaan yang dilakukan diatas deidapatkanlah R² tertinggi pada *kernel Gaussian* atau *Radial* dengan R² sebesar 58.53 % , sedangkan pada *kernel Linear* menghasilkan R² sebesar 51.43 % , dan R² pada *kernel Polynomial* merupakan akurasi terendah pada percobaan ini dengan nilai R² sebesar 36.03%.

5. KESIMPULAN

Dari penelitian yang dilakukan didapatkanlah hasil dan dapat diambil kesimpulan sebagai berikut:

- Percobaan pertama dengan melakukan *feature selection* dengan percobaan menggunakan matriks jarak didapatkan R² sebesar 27.47 % untuk *kernel Linear*, 35.70 % untuk *kernel Gaussian* dan 18.94 % untuk *kernel Polynomial*. Sedangkan pada percobaan tanpa menggunakan matriks jarak didapatkan R² sebesar 29.09 % untuk *kernel Linear*, 39.83 % untuk *kernel Gaussian* dan 20.93 % untuk *kernel Polynomial*.
- Percobaan kedua dengan tanpa melakukan *feature selection* dengan percobaan menggunakan matriks jarak didapatkan R² sebesar 7.22 % untuk *kernel Linear*, 47.67 % untuk *kernel Gaussian* dan 33.28 % untuk *kernel Polynomial*. Sedangkan pada percobaan tanpa menggunakan matriks jarak didapatkan R² sebesar 51.43 % untuk *kernel Linear*, 58.53 % untuk *kernel Gaussian* dan 36.03 % untuk *kernel Polynomial*.
- Dengan data yang digunakan tersebut didapatkanlah R² terbesar pada *Kernel Gaussian* dengan tanpa menggunakan *Feature Selection* dan tanpa matriks jarak.

DAFTAR PUSTAKA

- [1] World Health Organization, “**Global Tuberculosis Report**”, 2015.
Available at : http://www.who.int/tb/publications/global_report/en/
- [2] Indriani, D., Adiningsih, S., Mahmudiono, T, “**Faktor resiko yang mempengaruhi kejadian TB paru pada anak jalanan dengan studi kasus di Yayasan Insani Surabaya**”, Jurnal FKM UA, Surabaya, 2005.
- [3] Kementerian Kesehatan RI, “**Penderita penyakit Tuberkulosis di Dunia**”, Jakarta: Kemenkes RI, 2017.
- [4] Siagian, R. Y., “**Klasifikasi Parket Kayu Jati Menggunakan Metode Support Vector Machine (SVM)**”, Skripsi, Jurusan Teknik Informatika, Fakultas Teknologi Industri, Universitas, Jawa Barat, 2011.
- [5] S. Ansari and U. Sutar, “**Devanagari Handwritten Character Recognition using Hybrid Features Extraction and Feed Forward Neural Network Classifier (FFNN)**”, Int. J. Comput. Appl., vol. 129, no. 7, pp. 22–27, 2015.
- [6] Lumbanraja, F.R., Sani, R.M.S., Kurniawan, D., Irawati, A.R., “**Implementasi Metode Support Vector Machine Dalam Prediksi Persebaran Demam Berdarah Di Kota Bandar Lampung**”, Jurnal Komputasi, Vol. 7., No. 2, pp.63-73, 2019.
- [7] S.Shalev-Shwartz and Shai Ben-David, “**Understanding Machine Learning : From Theory to Algorithms**”, Cambridge University Press, 2014.
- [8] C. Cortes and V. Vapnik, “**SupportVector Networks**”, vol. 297, pp. 273–297, 1995.
- [9]Haslinda, J. M., “**Pengaruh Perencanaan Anggaran Dan Evaluasi Anggaran Terhadap Kinerja Organisasi Dengan Standar Biaya Sebagai Variabel Moderating Pada Pemerintah Daerah Kabupaten Wajo**” Jurnal Ilmiah Akuntansi Peradaban, 2016.
- [10]H. Bhavsar and M. H. Panchal, “**A Review on Support Vector Machine for Data Classification**”, Int. J. Adv. Res. Comput. Eng. Technol., vol. 1, no. 10, pp. 2278–1323, 2012.
- [11]Reitermanova, Z., “**Data Splitting. WDS’10 Proceedings of Contributed Papers. Czech Republic**”, Part 1:31-36, 2010.
- [12]D. Anguita, L. Ghelardoni, A. Ghio, L. Oneto, and S. Ridella, “**The ‘K’ in K-fold cross validation**,” ESANN 2012 proceedings, 20th Eur. Symp. Artif. Neural Networks, Comput. Intell. Mach. Learn., no. April, pp. 441–446, 2012.