

KLASIFIKASI KABUPATEN TERTINGGAL DI KAWASAN TIMUR INDONESIA DENGAN SUPPORT VECTOR MACHINE

Esa A Sari¹, Meilani T Br Saragih², Ilhamsyah A Shariati³, Sabiq Sofyan⁴, Rakai A Baihaqi⁵, Rani Nooraeni⁶

^{1,2,3,4,5,6}Politeknik Statistika STIS, Jalan Otto Iskandardinara No. 64C, Jakarta Timur, 13330, Indonesia
Email: ¹esasari16@gmail.com, ²meilanithereza28@gmail.com, ³aliilham61@gmail.com,
⁴sabiqsofyan98@gmail.com, ⁵sonar808@gmail.com, ⁶raninoor@stis.ac.id

(Naskah masuk: 4 November 2020, diterima untuk diterbitkan: 24 November 2020)

Abstrak

Kesenjangan pembangunan antar wilayah tampak dari adanya istilah daerah maju dan daerah tertinggal. Pengentasan daerah tertinggal masih menjadi persoalan khususnya di Kawasan Timur Indonesia (KTI). Dari 122 kabupaten yang ditetapkan sebagai daerah tertinggal pada RPJMN 2015-2019, terdapat 54 kabupaten yang menjadi daerah prioritas. Mayoritas wilayah yang terkategori sebagai daerah prioritas ini berada di KTI. Adanya perbedaan karakteristik pada setiap wilayah Indonesia membuat perlunya analisis tersendiri. Salah satu perbedaan karakteristik wilayah dapat dilihat berdasarkan indikator ketertinggalan daerah. Pada Kawasan Barat Indonesia (KBI), aspek ketertinggalan wilayah berfokus pada infrastruktur. Sementara, aspek ekonomi dan Sumber Daya Manusia (SDM) merupakan fokus aspek ketertinggalan wilayah di KTI. Tujuan penelitian ini ingin melihat hasil pengklasifikasian kawasan timur berdasarkan indikator ketertinggalan daerah tersebut yang mencakup indikator perekonomian, SDM, Kemampuan Keuangan Daerah, infrastruktur, aksesibilitas, dan karakteristik daerah menggunakan model supervised learning yaitu *Support Vector Machine* (SVM). Dasar pemilihan model pengklasifikasian dengan SVM adalah karena keunggulannya yang fleksible mencakup model regresi linier maupun *non-linear*. Hasil dari penelitian ini secara umum di Kawasan Timur Indonesia, kabupaten tertinggal memiliki level pembangunan yang lebih rendah dibandingkan kabupaten tidak tertinggal. Dengan menggunakan metode SVM dan fungsi *kernel* terbaiknya yaitu fungsi *kernel linear*, kabupaten di KTI mampu diklasifikasikan secara benar sebesar 87,23%. Berdasarkan kurva ROC diperoleh nilai AUC sebesar 0,862 yang artinya kinerja model sudah bagus. Sehingga metode SVM dengan *kernel linear* dapat menjadi salah satu rekomendasi dalam pengklasifikasian kabupaten tertinggal atau tidak tertinggal

Kata kunci: *Kabupaten Tertinggal, Kawasan Timur Indonesia (KTI), Support Vector Machine*

CLASSIFICATION OF UNDERDEVELOPED DISTRICTS IN EASTERN INDONESIA WITH SUPPORT VECTOR MACHINE

Abstract

The development gap between regions can be seen from the terms developed and disadvantaged regions. Alleviating disadvantaged areas is still a problem, especially in Eastern Indonesia. Of the 122 districts designated as underdeveloped areas in the 2015-2019 RPJMN, there are 54 districts that are priority areas. The majority of areas categorized as priority areas are located in Eastern Indonesia. The different characteristics of each region of Indonesia require a separate analysis. One of the differences in regional characteristics can be seen based on the regional backwardness indicators. In the Western Region of Indonesia, the backward aspect of the region focuses on infrastructure. Meanwhile, the economic and Human Resources (HR) aspects are the focus of the underdeveloped aspects of the region in Eastern Indonesia. The purpose of this study is to see the results of the classification of the eastern region based on the regional underdevelopment indicators which include economic, human resources, regional financial capabilities, infrastructure, accessibility, and regional characteristics using a supervised learning model, namely Support Vector Machine (SVM). The basis for selecting a classification model with SVM is because of its flexibility to include both linear and non-linear regression models. The results of this study generally in Eastern Indonesia, underdeveloped districts have a lower level of development than non-lagging districts. By using the SVM method and the best kernel function, namely the linear kernel function, districts in Eastern Indonesia were able to be classified correctly by 87.23%. Based on the ROC curve, the AUC value is 0.862, which means that the model's performance is good. So that the SVM method with a linear kernel can be one of the recommendations in the classification of underdeveloped or not underdeveloped district.

Keywords: *Underdeveloped Districts, Eastern Indonesia, Support Vector Machine*

1. PENDAHULUAN

Kesenjangan pembangunan antar wilayah merupakan salah satu permasalahan pembangunan yang sering terjadi di negara berkembang, tak terkecuali Indonesia. Akibat yang dapat ditimbulkan dari kesenjangan tersebut adalah munculnya wilayah yang dikategorikan sebagai daerah maju dan daerah tertinggal. Daerah tertinggal adalah daerah kabupaten yang wilayah serta masyarakatnya kurang berkembang dibandingkan dengan daerah lain dalam skala nasional (Perpres RI No. 131 Tahun 2015). Ketertinggalan suatu daerah dapat disebabkan karena kurang berkembangnya kegiatan ekonomi masyarakat, rendahnya kualitas sumber daya manusia, tidak tersedianya sarana dan prasarana yang memadai, meningkatnya kejadian bencana alam dan konflik, serta rendahnya akses transportasi, telekomunikasi dan informasi.

Sejak tahun 2007, daerah tertinggal sudah menjadi target utama program percepatan pembangunan. Pemerintah sudah melakukan banyak upaya untuk mengatasi adanya daerah tertinggal, salah satunya Pembangunan Sosial Ekonomi Daerah Tertinggal (P2SEDT). Upaya ini cukup berhasil, terbukti pada 2010 jumlah daerah tertinggal sebanyak 183 kabupaten berkurang menjadi 122 kabupaten pada 2015. Dari total 122 kabupaten yang ditetapkan sebagai daerah tertinggal pada RPJMN 2015-2019, terdapat 54 kabupaten yang menjadi daerah prioritas. Mayoritas wilayah yang terkategori sebagai daerah prioritas ini merupakan kabupaten yang ada di Kawasan Timur Indonesia (KTI). Hal ini membuktikan perlunya perhatian lebih terhadap pembangunan daerah di KTI.

Pemerataan pembangunan perlu dilakukan demi mencegah munculnya daerah tertinggal, karena masalah ketertinggalan pembangunan suatu daerah dapat meluas menjadi masalah sosial dan ekonomi masyarakat. Untuk itu, perlu dilakukan pengklasifikasian wilayah agar dapat menentukan prioritas dalam mencapai pemerataan pembangunan secara cepat dan tepat sasaran. Menurut [1] diperlukan klasifikasi yang berbeda untuk setiap kawasan wilayah di Indonesia. Hal ini disebabkan setiap kawasan wilayah memiliki karakteristik yang berbeda. Salah satu perbedaan karakteristik wilayah dapat dilihat melalui aspek utama ketertinggalan antar kawasan. Pada Kawasan Barat Indonesia (KBI), aspek ketertinggalan wilayah berfokus pada infrastruktur. Sementara, aspek ekonomi dan Sumber Daya Manusia (SDM) merupakan fokus aspek ketertinggalan wilayah di KTI.

Penentuan atau pengklasifikasian daerah tertinggal menurut Perpres RI No. 131 Tahun 2015 dilakukan dengan indeks komposit berdasarkan 6

kriteria dan 27 sub indikator ketertinggalan. Kriteria yang digunakan yaitu perekonomian masyarakat, sumber daya manusia, sarana dan prasarana, kemampuan keuangan daerah, aksesibilitas, dan karakteristik daerah. Penggunaan indeks komposit dalam proses klasifikasi daerah tertinggal di Indonesia masih belum optimal. Hal ini disebabkan masih ada daerah-daerah yang tidak termasuk kategori tertinggal tetapi cenderung memiliki karakteristik yang sama dengan daerah tertinggal. Oleh karena itu, perlu adanya penggunaan metode lain yang lebih baik dalam melakukan klasifikasi objek, dalam hal ini wilayah kabupaten di Indonesia.

Berbagai metode statistik dapat diterapkan dalam melakukan analisis klasifikasi. Salah satunya adalah *Support Vector Machine* (SVM). *Support Vector Machine* (SVM) adalah suatu teknik yang dapat memisahkan dua himpunan data dari dua kelas yang berbeda dengan memaksimalkan batas fungsi pemisah (hyperplane) [2]. Salah satu kelebihan dari metode ini adalah dapat melakukan pengklasifikasian dan mengatasi regresi dengan *linear* maupun *non-linear*. SVM memiliki tingkat akurasi klasifikasi yang lebih baik dibandingkan dengan metode klasifikasi lainnya, seperti *Naïve Bayes*, *C.45*, dan *KNN* [2].

Beberapa penelitian telah dilakukan dengan memilih metode klasifikasi terbaik, yakni dilakukan oleh [3] terhadap klasifikasi wilayah desa-perdesaan dan desa-perkotaan wilayah Kabupaten Semarang menggunakan metode *Support Vector Machine*. diperoleh hasil rata-rata akurasi sebesar 90%. Kemudian penelitian yang kedua yakni dilakukan [4] mengenai Klasifikasi Wilayah Demam Berdarah dengan Metode SVM di Kota Medan. Hasil yang diperoleh pada penelitian tersebut adalah proses klasifikasi wilayah demam berdarah memperoleh akurasi pengujian sebanyak 75%.

Berdasarkan latar belakang yang telah diuraikan, maka penelitian ini akan menggambarkan karakteristik daerah di Kawasan Timur Indonesia, memprediksi status ketertinggalan kabupaten di Kawasan Timur Indonesia berdasarkan beberapa indikator ketertinggalan daerah, dan menentukan ketepatan klasifikasi kabupaten tertinggal dengan *Support Vector Machine* (SVM).

2. METODE PENELITIAN

Unit analisis yang akan diteliti dalam penelitian ini adalah seluruh kabupaten/kota di Kawasan Timur Indonesia yaitu kabupaten/kota yang ada di pulau Sulawesi, kepulauan Nusa Tenggara (kecuali Bali), kepulauan Maluku, dan pulau Papua. provinsi Nusa Tenggara Barat, Nusa Tenggara Timur, Sulawesi Utara, Gorontalo, Sulawesi Tengah, Sulawesi Barat,

Sulawesi Selatan, Sulawesi Tenggara, Maluku Utara, Maluku, Papua Barat, dan Papua.

Data yang digunakan dalam penelitian ini merupakan data sekunder yang bersumber dari Badan Pusat Statistik. Adapun variabel-variabel yang digunakan adalah persentase rumah tangga pengguna air bersih, persentase rumah tangga pengguna listrik, persentase desa yang mempunyai konflik satu tahun terakhir, persentase penduduk miskin, angka harapan hidup, rata-rata lama sekolah, kemampuan keuangan daerah, jumlah puskesmas, persentase pengeluaran makanan, persentase pekerja di sektor pertanian, persentase desa terjangkau internet, dan status daerah.

2.1. Supervised Learning

Supervised learning atau analisis klasifikasi merupakan proses mencari suatu model atau fungsi yang mampu menjelaskan dan membedakan data ke dalam kelompok yang telah memiliki target/label/class [5]. Analisis klasifikasi juga dapat melakukan prediksi dari suatu kelas saat tidak diketahui label dari kelas tersebut. Dalam melakukan analisis klasifikasi, data dibagi menjadi dua bagian yaitu *training set* untuk membuat model dan *test set* untuk menilai keakuratan klasifikasi yang telah dilakukan [6]. Salah satu metode yang dapat digunakan untuk membagi dataset sebagai *training set* dan *test set* adalah *k-fold cross validation*. Metode ini membagi data secara random menjadi beberapa (k) bagian kemudian dilakukan *training* dengan menggunakan beberapa bagian data dan *testing* dengan bagian lainnya. Jumlah k yang biasa digunakan adalah 10 (*10-fold cross validation*). Metode *10-fold cross validation* sering digunakan karena cenderung memberikan estimasi akurasi yang tidak bias dibandingkan dengan metode lainnya seperti *split validation*, *leave-one-out cross validation* dan *bootstrap*.

2.2. Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi – fungsi linier dalam sebuah fitur yang berdimensi tinggi dan dilatih dengan menggunakan algoritma pembelajaran yang didasarkan pada teori optimasi. Cara kerja SVM adalah menemukan *hyperplane* yang baik sebagai pemisah dua kelas yang berbeda dengan memaksimalkan margin, dimana margin maksimum diketahui melalui *Support Vector* [7].

SVM juga merupakan metode yang dapat diterapkan pada masalah *non-linear*. Dengan menggunakan metode *kernel trick* yang mencari *hyperplane* dengan cara mentransformasi dataset ke ruang vektor yang berdimensi lebih tinggi (*feature space*), kemudian proses klasifikasi dilakukan pada

feature space tersebut [8]. Sehingga fungsi yang dihasilkan dari pelatihan yaitu:

$$f(x_d) = \sum_{i=1}^{ns} \alpha_i y_i K(x_i, x_d) + b \quad (1)$$

dimana :

ns: jumlah *Support Vector*

α : alpha, Pengali *Lagrange*

b : bias

y : Label / kelas dari data

r : *coefficient*

p : pangkat atau derajat polinomial

x : *Support vector*

Beberapa fungsi *kernel* yang umum digunakan pada *Support Vector Machine* (SVM) adalah fungsi *kernel Linear*, *Radial Basis Function* (RBF), dan *Polynomial*. Fungsi *kernel* dan parameter yang digunakan dalam analisis SVM sangat memengaruhi tingkat akurasi yang akan dihasilkan. [9].

2.3. Evaluasi Kinerja Klasifikasi

Pengukuran kinerja klasifikasi dilakukan untuk mengevaluasi kinerja model klasifikasi dalam memprediksikan kelas suatu data [10]. Suatu model klasifikasi ditetapkan sebagai model terbaik berdasarkan kriteria akurasi, *specificity*, *sensitivity*, *precision*, *recall*, dan *F Measure* yang tertinggi. Hasil dari jumlah observasi yang benar diklasifikasikan ataupun yang misklasifikasi dapat disusun dalam sebuah *confusion matrix* seperti pada Tabel 1.

Tabel 1. *Confusion Matrix*

Kelas Aktual	Kelas Prediksi	
	Positif	Negatif
Positif	TP	FN
Negatif	FP	TN

Keterangan:

TP: Banyaknya prediksi benar pada kelas positif

FP: Banyaknya prediksi salah pada kelas positif

TN: Banyaknya prediksi benar pada kelas negatif

FN: Banyaknya prediksi salah pada kelas negatif

Berdasarkan tabel *confusion matrix* dapat dihitung nilai akurasi, *specificity*, *sensitivity*, *precision*, *recall*, dan *F Measure* adalah sebagai berikut:

$$\text{Akurasi} = \frac{TN+TP}{TN+TP+FN+FP}$$

$$\text{Sensitivity} = \frac{TP+FN}{TN}$$

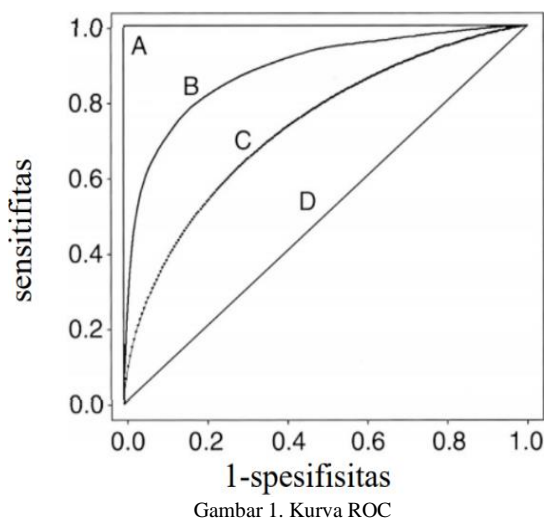
$$\text{Specificity} = \frac{TN+FP}{TP}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Recall} = \frac{TP}{TP+FN}$$

$$\text{F measure} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

Secara visual model-model klasifikasi yang telah dibentuk dapat dibandingkan dengan menggunakan kurva ROC (*Receiver Operating Characteristics*). Sumbu vertikal pada kurva ROC menggambarkan *true positive rate* (sensitivitas), sedangkan sumbu horizontal menggambarkan *false positive* (1-spesifisitas). Luas area di bawah kurva ROC atau yang disebut dengan AUC (*Area Under Cover*) menunjukkan akurasi model, dimana semakin jauh kurva dengan garis diagonal, maka semakin akurat model klasifikasi. Model klasifikasi dengan akurasi sempurna akan memiliki luas area sebesar 1,0 dan akurasi terburuk memiliki luas area 0,5. Sehingga berdasarkan Gambar 1, kurva A memiliki performa yang lebih baik dibandingkan D.



Kategori pengklasifikasian model berdasarkan nilai AUC [11] diringkas pada Tabel 2.

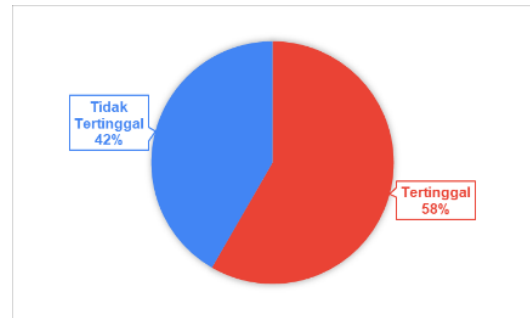
Tabel 2. Nilai AUC dan Kategori Hasil Klasifikasi

Nilai AUC	Kategori Hasil Klasifikasi
0,9 - 1,0	Excellent
0,8 - 0,9	Good
0,7 - 0,8	Fair
0,6 - 0,7	Poor
0,5 - 0,6	Fail

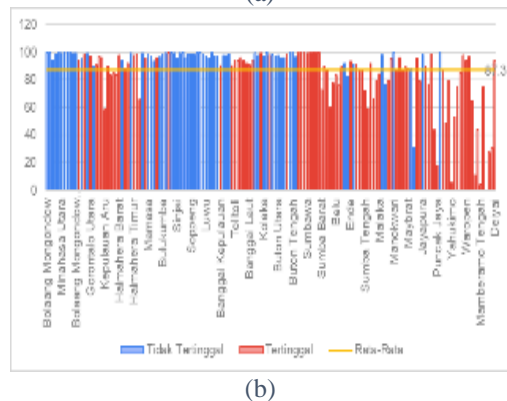
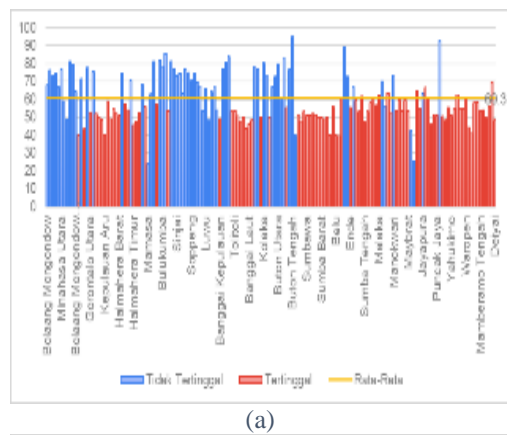
3. HASIL DAN PEMBAHASAN

3.1. Karakteristik Daerah Tertinggal

Kawasan Timur Indonesia (KTI) memiliki 156 kabupaten berasal dari 12 provinsi. Sebesar 58% dari total 156 kabupaten di KTI berstatus sebagai daerah tertinggal.

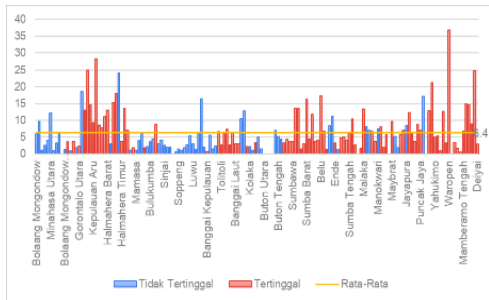


Berdasarkan Gambar 2. proporsi kabupaten tertinggal lebih banyak dibandingkan dengan kelas kabupaten tidak tertinggal.



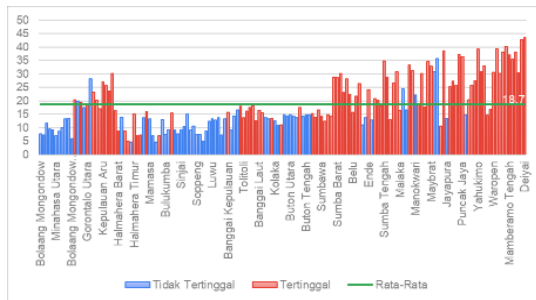
Gambar 3. Kriteria Infrastruktur (a) Persentase Ruta Pengguna Air Bersih (b) Persentase Ruta Pengguna Listrik

Gambar 3. menunjukkan grafik batang kriteria infrastruktur yang terdiri dari variabel persentase rumah tangga pengguna air bersih dan listrik. Grafik batang yang berwarna merah merupakan kabupaten tertinggal yang menunjukkan bahwa sebagian besar kabupaten tertinggal memiliki persentase rumah tangga pengguna air bersih di bawah rata-rata KTI. Sementara itu, untuk persentase rumah tangga pengguna listrik, sebagian besar kabupaten dengan kelas tidak tertinggal rumah tangganya sudah menggunakan listrik. Sehingga persentase rumah tangga pengguna listrik kabupaten dengan kelas tidak tertinggal sudah berada di atas rata-rata KTI.



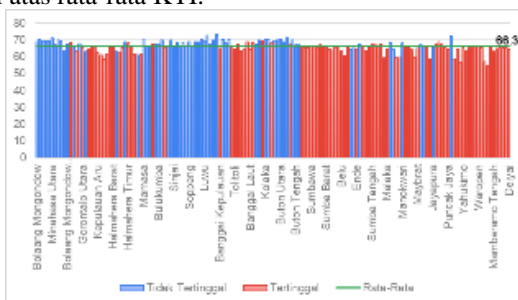
Gambar 4. Persentase Desa yang Mempunyai Konflik Satu Tahun Terakhir

Pada gambar 4 di atas ditunjukkan persentase desa yang mempunyai konflik satu tahun terakhir dengan grafik batang berwarna merah sebagai kabupaten yang tertinggal. Menurut grafik batang tersebut sebagian besar kabupaten dengan kelas tertinggal, persentase desa mempunyai konflik satu tahun terakhir berada di atas rata-rata KTI.

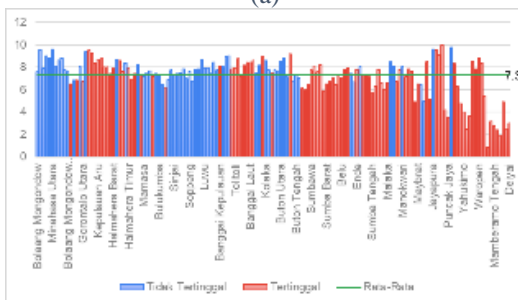


Gambar 5. Persentase Penduduk Miskin

Rata-rata persentase kemiskinan di KTI adalah 18,7%. Grafik batang yang berwarna merah menunjukkan bahwa sebagian besar kabupaten tertinggal memiliki rata-rata persentase kemiskinan di atas rata-rata KTI.



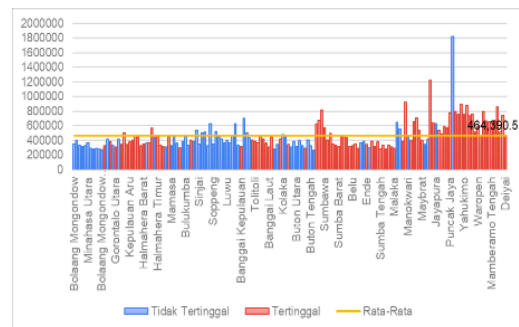
(a)



(b)

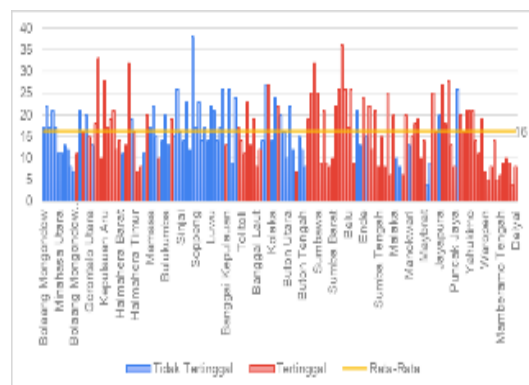
Gambar 6. Kriteria SDM (a) Angka Harapan Hidup (b) Rata-Rata Lama Sekolah

Gambar 6. menunjukkan grafik batang kriteria sumber daya manusia yang terdiri dari variabel angka harapan hidup dan rata-rata lama sekolah. Di Kawasan Timur Indonesia rata-rata angka harapan hidupnya adalah 66,8 tahun. Sebagian besar kabupaten dengan kelas tertinggal memiliki angka harapan hidup di bawah rata-rata KTI. Pada kelas kabupaten tertinggal terdapat kabupaten dengan rata-rata lama sekolah terendah yaitu Kabupaten Nduga, Papua dengan rata-rata lama sekolah sebesar 0,85 tahun. Selain itu, sebanyak 21 Kabupaten lainnya memiliki rata-rata lama sekolah kurang dari 6 tahun sehingga dapat disimpulkan rata-rata penduduk yang berumur 25 tahun ke atas hanya bersekolah kurang dari 6 tahun atau tidak tamat pendidikan sekolah dasar. Secara keseluruhan, rata-rata lama sekolah di Kawasan Timur Indonesia adalah 7,3 tahun.

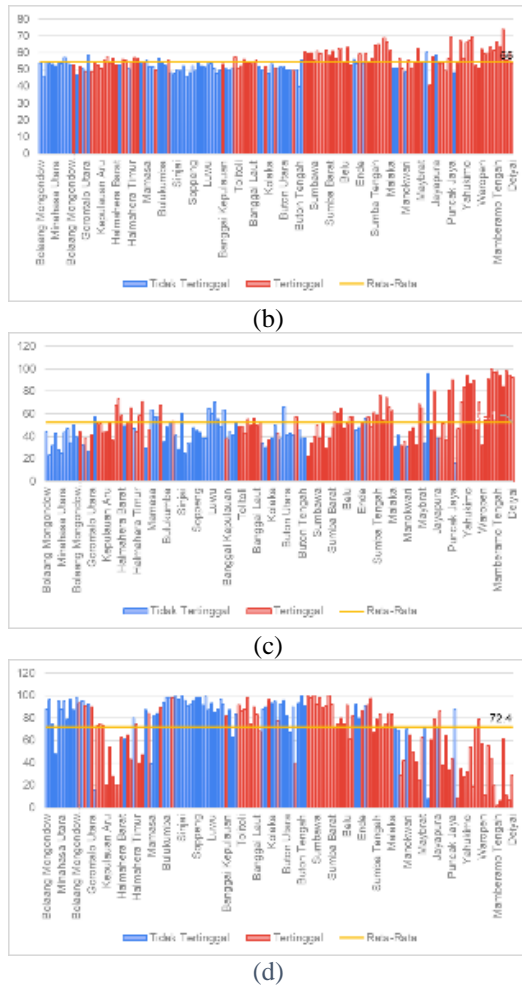


Gambar 7. Kemampuan Keuangan Daerah

Gambar 7. menunjukkan kemampuan keuangan daerah di kabupaten-kabupaten KTI. Grafik batang yang berwarna merah mewakili kabupaten tertinggal. Rata-rata kemampuan keuangan daerah KTI adalah Rp464.390,5. Menurut grafik batang di atas jumlah kabupaten dengan kemampuan keuangan daerah di bawah rata-rata KTI didominasi oleh kabupaten dengan kelas tidak tertinggal. Sementara itu, kabupaten dengan kelas tertinggal memiliki kemampuan keuangan daerah di atas rata-rata KTI terutama kabupaten-kabupaten di Pulau Papua.



(a)



Gambar 8. Variabel Tambahan (a) Jumlah Puskesmas (b) Persentase Pengeluaran Makanan (c) Persentase Pekerja di Sektor Pertanian (d) Persentase Desa Terjangkau Internet

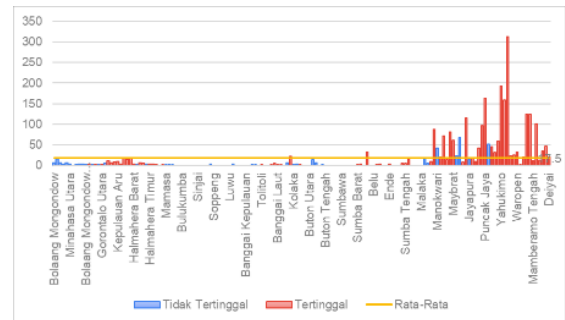
Gambar 8. menunjukkan variabel-variabel tambahan yang masih berkaitan dengan klasifikasi tertinggal atau tidaknya sebuah daerah yaitu jumlah puskesmas, persentase pengeluaran makanan, persentase pekerja di sektor pertanian, dan persentase desa terjangkau internet. Grafik batang berwarna merah menunjukkan kabupaten dengan kelas tertinggal.

KTI memiliki rata-rata jumlah puskesmas sebanyak 16 unit. Baik kabupaten dengan kelas tertinggal maupun tidak tertinggal, sebagian besar memiliki jumlah puskesmas di bawah rata-rata jumlah puskesmas KTI. Jumlah puskesmas terbanyak berada di Kabupaten Bone, Sulawesi Selatan dengan jumlah puskesmas sebanyak 38 unit.

Sebagian besar persentase pengeluaran makanan kabupaten yang berada di atas rata-rata adalah kabupaten dengan kelas tertinggal, bahkan sebagai salah satu kelas kabupaten tertinggal, Kabupaten Puncak memiliki persentase pengeluaran makanan tertinggi yaitu sebesar 73,68 persen.

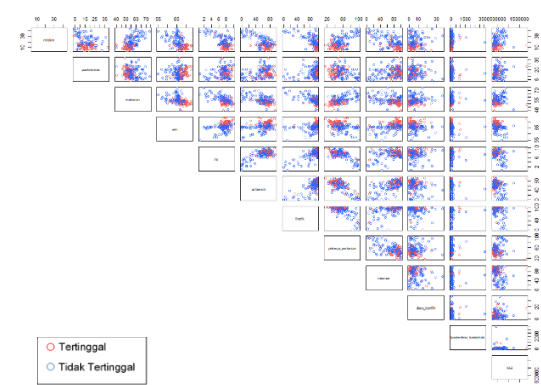
Sebagian besar kabupaten tertinggal masih mengandalkan sektor pertanian, oleh karena itu kabupaten tertinggal memiliki persentase pekerja di

sektor pertanian yang tergolong tinggi, terlebih lagi di Provinsi Papua jika dilihat berdasarkan grafik. Masih cukup banyak kabupaten dengan kelas tertinggal memiliki persentase desa terjangkau internet yang rendah. Provinsi Papua dengan topografinya bergunung-gunung yang membuat sulit menjangkau sinyal internet menunjukkan sebagian besar wilayah kabupatennya memiliki persentase desa terjangkau internet yang tergolong rendah.



Gambar 9. Jumlah Desa dengan Jarak Pelayanan Kesehatan Lebih Dari 5 Km

Gambar 9. menunjukkan jumlah desa dengan jarak pelayanan kesehatan lebih dari 5 km. grafik batang berwarna merah menunjukkan kabupaten dengan kelas tertinggal. Sebagian besar kabupaten dengan kelas tertinggal memiliki jumlah desa dengan jarak pelayanan kesehatan lebih dari 5 km di atas rata-rata KTI. Kabupaten-kabupaten tersebut sebagian besar berada di Pulau Papua. Hal ini menunjukkan aksesibilitas terhadap pelayanan kesehatan di KTI terutama di Pulau Papua masih sulit diraih.



Gambar 10. Pola Persebaran Data Seluruh Variabel

Pada Gambar 10. menunjukkan bahwa terdapat beberapa pola data menyebar yang sulit dipisahkan dengan fungsi pemisah linier. Sehingga diperlukan bantuan *kernel* untuk mendapatkan fungsi pemisah yang optimal.

3.2. Parameter Terbaik untuk *Hyperplane* pada setiap Fungsi *Kernel*

Pada penelitian ini menggunakan tiga fungsi *kernel* yaitu fungsi *kernel Linear*, *Polynomial*, dan *Radial Basis Function (RBF)* pada fungsi pemisah

(*hyperplane*) SVM. Dalam menentukan parameter terbaik pada setiap fungsi *kernel* digunakan metode *10-fold cross validation*. Setelah dilakukan seleksi parameter, diperoleh parameter terbaik untuk *hyperplane* dengan fungsi *kernel linear* $C = 1$, fungsi *kernel polynomial* $p = 2$ dan $C = 0,25$, dan fungsi *kernel RBF* $\sigma = 0,1222638$ dan $C = 1$.

3.3. Evaluasi Kinerja Klasifikasi

Setelah diperoleh parameter terbaik pada setiap fungsi *kernel*, langkah selanjutnya yaitu menentukan fungsi *kernel* terbaik pada pengklasifikasian kabupaten tertinggal atau tidak tertinggal di Kawasan Timur Indonesia. Dengan menggunakan *confusion matrix* dapat dinilai seberapa baik sebuah *classifier*. Tabel *confusion matrix* merupakan hasil pengaplikasian model pada data *testing*. Dari *confusion matrix* diperoleh berbagai *metric* evaluasi *classifier* yaitu akurasi, *specificity*, *sensitivity*, *precision*, *recall*, dan *F Measure*. Selanjutnya, dibandingkan hasil klasifikasi setiap fungsi *kernel* untuk menentukan fungsi *kernel* terbaik. Perbandingan *metric* evaluasi *classifier* dapat ditunjukkan pada Tabel 3.

Tabel 3. Perbandingan *Metric* Evaluasi *Classifier*

	<i>Linier</i>	<i>Polynomial</i>	RBF
Akurasi	0,8723	0,8511	0,8511
<i>sensitivity</i>	0,8235	0,8235	0,8235
<i>Specificity</i>	0,9000	0,8667	0,8667
<i>Precision</i>	0,8235	0,7778	0,7778
<i>Recall</i>	0,8235	0,8235	0,8235
<i>F Measure</i>	0,8235	0,8	0,8

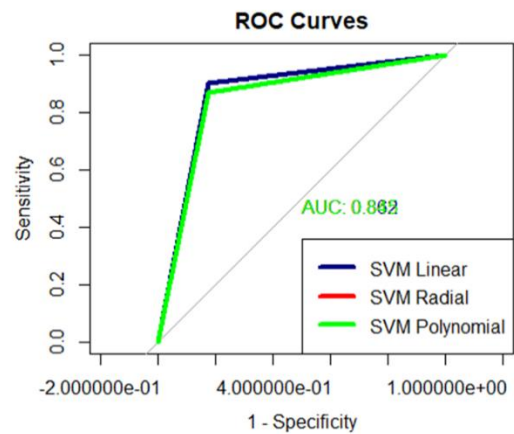
Berdasarkan perbandingan *metric* evaluasi *classifier*, fungsi *kernel linear* memiliki akurasi, *specificity*, *sensitivity*, *precision*, *recall*, dan *F Measure* yang lebih tinggi dibandingkan fungsi *kernel* lainnya. Sehingga fungsi *kernel linear* menjadi fungsi *kernel* terbaik dengan parameter $C = 1$. Hasil pengklasifikasian menggunakan metode SVM dengan *kernel linear* dapat dilihat pada Tabel 4.

Tabel 4. Tabel klasifikasi dari model SVM *kernel Linear*

Aktual	Prediksi	
	Tidak Tertinggal	Tertinggal
Tidak Tertinggal	14	3
Tertinggal	3	27

Berdasarkan Tabel 4. diperoleh akurasi model sebesar 87,23%. Artinya, model SVM dengan *kernel linear* mampu mengklasifikasikan kabupaten baru sebagai tertinggal atau tidak tertinggal dengan tepat sebesar 87,23%. Sedangkan *error rate* model sebesar 12,77%, dimana kabupaten yang awalnya dikategorikan tidak tertinggal menjadi tertinggal ada 3 yaitu Kabupaten Halmahera Tengah, Kabupaten Kaimana, dan Kabupaten Jayapura. Sementara

kabupaten yang awalnya dikategorikan tertinggal menjadi tidak tertinggal juga ada 3 yaitu Kabupaten Mamuju Tengah, Kabupaten Sigi, dan Kabupaten Bombana.



Gambar 11. Kurva ROC dari evaluasi model SVM

Berdasarkan kurva ROC yang terbentuk pada Gambar 11. menunjukkan ukuran akurasi model dengan luasan di bawah kurva (AUC) sebesar 0,862 yang artinya masuk ke dalam kategori "Good", sehingga model yang terbentuk kinerjanya sudah bagus.

4. KESIMPULAN

Secara umum, di Kawasan Timur Indonesia kabupaten tertinggal memiliki level pembangunan yang lebih rendah dibandingkan kabupaten tidak tertinggal. Dengan menggunakan metode SVM diperoleh fungsi *kernel* terbaik yaitu fungsi *kernel linear* dengan parameter $C=1$ dan mampu mengklasifikasikan secara benar sebesar 87,23%. Sedangkan *error rate* model sebesar 12,77%, dimana kabupaten yang awalnya dikategorikan tidak tertinggal menjadi tertinggal ada 3 yaitu Kabupaten Halmahera Tengah, Kabupaten Kaimana, dan Kabupaten Jayapura. Sementara kabupaten yang awalnya dikategorikan tertinggal menjadi tidak tertinggal juga ada 3 yaitu Kabupaten Mamuju Tengah, Kabupaten Sigi, dan Kabupaten Bombana.

Berdasarkan kurva ROC diperoleh nilai AUC sebesar 0,862 yang artinya kinerja model sudah bagus. Selain itu, nilai akurasi, *sensitivity*, *specificity*, *precision*, *recall*, dan *F Measure* juga lebih dari 80%. Sehingga metode SVM dengan *kernel linear* dapat menjadi salah satu rekomendasi dalam pengklasifikasian kabupaten tertinggal atau tidak tertinggal.

5. DAFTAR PUSTAKA

- [1] S. I. Oktora. 2015. "Analisis Multivariate Adaptive Regression Splines (MARS) pada Prediksi Ketertinggalan Kabupaten Tahun 2014"., *Jurnal Aplikasi Statistika & Komputasi Statistik*, vol. 7(2), pp. 115-128

- [2] V. Vapnik dan C. Cortes. 1995. "Support-vector networks". *Machine learning*, vol. 20(3), pp. 273-297.
- [3] M. S. Sari, D. Safitri., dan S. Sugito. 2014. "Klasifikasi Wilayah Desa-Perdesaan dan Desa-Perkotaan Wilayah Kabupaten Semarang dengan Support Vector Machine (SVM)"., *Jurnal Gaussian*, vol.3(4), pp. 751-760
- [4] H. Situmorang. 2015. "Klasifikasi Wilayah Demam Berdarah dengan Metode SVM di Kota Medan". *Tesis*, Universitas Sumatera Utara, Indonesia.
- [5] J. Han dan M. Kamber. 2006. *Data Mining: Concept and Techniques Second Edition*. United State of America: Morgan Kaufman Publishers
- [6] R. Septiani, I. P. A. Citra., dan A. S. A. Nugraha. 2019. "Perbandingan Metode Supervised Classification dan Unsupervised Classification terhadap Penutup Lahan di Kabupaten Buleleng"., *Jurnal Geografi*, vol. 16(2), pp. 91-96.
- [7] J.P. Maulana dan Irhamah. 2018. "Klasifikasi Kabupaten di Provinsi Jawa Timur Berdasarkan Indikator Daerah Tertinggal dengan metode Support Vector Machine (SVM) dan Entropy Based Fuzzy Support Vector Machine (EFSVM)"., *Inferensi*, vol. 1(1), pp. 9-15
- [8] M. Affandes. 2015. "Penerapan Metode Support Vector Machine (SVM) Menggunakan Kernel Radial Basis Faunction (RBF) Pada Klasifikasi Tweet"., *Jurnal Sains dan Teknologi Industri*, vol. 12(2), pp. 189-197
- [9] N. R. Feta dan A. R. Ginanjar. 2019. "Komparasi Fungsi Kernel Metode Support Vector Machine Untuk Pemodelan Klasifikasi Terhadap Penyakit Tanaman Kedelai"., *BRITech, Jurnal Ilmiah Ilmu Komputer, Sains dan Teknologi Terapan*, vol.1(1), pp. 33-39
- [10] A. H. Baksir., A. Fuad., F. Tempola., R. Rosihan. 2020. "Prediksi Tingkat Kualitas Kesuburan Pria Dengan Jaringan Saraf Tiruan Backpropagation". *JIKO (Jurnal Informatika dan Komputer)*. Vol. 3 no. 2, pp. 107-112.
- [11] F. Gorunescu, 2011. "Data Mining: Concepts, Models and Techniques". New York: Springer-Verlag.