

**ANALISIS SENTIMEN ALIH FUNGSI  
LAHAN PERTANIAN INDONESIA 2017 - 2021**

**Alvina Maharani Hasibuan, Matthew Martianus Henry, M. Iqbal**

**Anaraunga Team  
14030103**



**INSTITUT PERTANIAN BOGOR  
2021**

## ABSTRAK

Perkembangan infrastruktur merupakan tolak ukur kemajuan suatu negara di dunia. Tak terkecuali di Indonesia, perkembangan infrastruktur yang dilakukan pemerintah seringkali melibatkan alih fungsi lahan pertanian. Alih fungsi lahan pertanian untuk pembangunan dan pengembangan infrastruktur kerap kali menimbulkan pro dan kontra di masyarakat. Tujuan penelitian ini adalah membuat suatu model analisis sentimen kebijakan alih fungsi lahan pertanian di Indonesia. Metode yang digunakan untuk analisis sentimen adalah model *logistic regression* dan *support vector machine*. Hasil penelitian ini menunjukkan bahwa model *machine learning* terbaik untuk analisis sentimen alih fungsi lahan pertanian Indonesia adalah *support vector machine* dengan akurasi skor validasi sebesar 81%. Akurasi skor validasi model *logistic regression* adalah 80%. Model dengan akurasi skor validasi terbaik yaitu *support vector machine* digunakan pada data uji dan menghasilkan skor akurasi sebesar 82.2%.

Kata Kunci : alih fungsi lahan pertanian, *logistic regression*, *support vector machine*.

## PENDAHULUAN

### 1. Latar Belakang

Pada pasal 1 nomor 15 UU No. 41 Tahun 2009 disebutkan bahwa alih fungsi lahan pertanian adalah perubahan fungsi lahan pertanian pangan berkelanjutan menjadi bukan lahan pertanian pangan berkelanjutan baik secara tetap maupun sementara. Dalam hal ini alih fungsi lahan sering kali dilakukan untuk diubah menjadi industri, perumahan, maupun bangunan lainnya. Beberapa implementasi dari kebijakan alih fungsi lahan pertanian tidak sesuai dengan UU No. 41 Tahun 2009 yang mengatur tentang perlindungan lahan pertanian pangan berkelanjutan. Sebagaimana dilansir Kompas (2019), Menteri Pertanian RI, Syahrul Yasin Limpo membenarkan alih fungsi lahan pertanian ke lahan non pertanian masih berlangsung bahkan cenderung meningkat. Hal ini jika dibiarkan akan mengancam ketahanan pangan dan memaksa pemerintah untuk membuka keran impor untuk memenuhi kebutuhan pangan masyarakat luas. Alih fungsi lahan pertanian yang masif ini secara tidak langsung bertentangan dengan Pasal 3b UU No. 41 Tahun 2009 yang berbunyi “Menjamin tersedianya lahan pertanian pangan secara berkelanjutan”. Potongan ayat tersebut menunjukkan kewajiban pemerintah dan juga masyarakat dalam menjaga ketersediaan lahan pertanian sehingga menghindari keadaan rawan pangan melalui pengendalian alih fungsi lahan pertanian. Pengembangan infrastruktur negara yang kerap kali diikuti dengan alih fungsi lahan pertanian menuai pro dan kontra di masyarakat.

Penggunaan media sosial di Indonesia semakin berkembang setiap tahunnya. Platform media sosial seperti Facebook, Twitter, maupun Instagram menjadi tempat mengekspresikan pendapat dan ekspresi seseorang dalam berbagai topik. Salah satu media sosial yang diminati penduduk Indonesia adalah Twitter. Twitter sebagai salah satu platform media sosial memiliki fitur untuk memberikan cuitan (*tweet*) mengenai pendapat

seseorang. *Tweet* tersebut dibatasi hanya 140 karakter. Pengguna dapat secara bebas mengemukakan pendapatnya mengenai sesuatu pada platform Twitter, serta memberikan tanggapan terhadap *tweet* orang lain. Jumlah pengguna Twitter di Indonesia sebanyak 15.1 juta penduduk dan menduduki peringkat lima dunia pada April 2021 (Statista 2021). Pengguna Twitter di Indonesia sebagian besar adalah generasi muda dengan jumlah terbanyak berada pada rentang usia 25 hingga 34 tahun.

Analisis sentimen cuitan pada twitter dapat digunakan untuk mengetahui penilaian masyarakat terhadap kebijakan di bidang pertanian yang ditetapkan pemerintah. Penelitian sebelumnya mengenai *sentiment analysis* kebijakan pemerintah pada platform Twitter dilakukan oleh Nandini *et al* (2018) mengenai kebijakan impor beras pemerintah menggunakan model *support vector machine* dan pembobotan jumlah *retweet*. Penelitian lainnya yang serupa dilakukan pula oleh Dunnmon *et al* (2017) mengenai analisis sentimen komunitas pertanian di Amerika Serikat berkaitan dengan ketahanan pangan di Amerika Serikat. Dalam penelitian tersebut, Dunnmon *et al* (2017) mencoba menggunakan model *support vector machine*, *logistic regression* dan *dense neural network*.

## **2. Tujuan**

Penelitian dilakukan dengan tujuan sebagai berikut:

- a. Membuat model klasifikasi dengan algoritma *machine learning* untuk menemukan sentimen dari *tweet* yang berhubungan dengan kebijakan alih fungsi lahan pertanian.

## **3. Manfaat**

Penelitian ini diharapkan dapat:

- a. Mempermudah analisis kebijakan pemerintah yang melibatkan alih fungsi lahan pertanian di Indonesia melalui analisis sentimen pendapat masyarakat dengan model klasifikasi yang dibuat.

## **4. Batasan Penelitian**

Batasan penelitian ini meliputi:

- a. Data yang digunakan adalah *tweet* mengenai alih fungsi lahan pertanian di Twitter pada rentang waktu 2017 hingga Juli 2021.
- b. Data yang diambil hanya berupa opini pengguna Twitter dan tidak meliputi *tweet* dari *news portal*.
- c. Pengolahan data dilakukan menggunakan bahasa pemrograman Python 3 dengan GUI Jupyter Notebook dan Google Colab.

## STUDI LITERATUR

### Logistic Regression

*Logistic regression* (LR) merupakan metode analisis yang digunakan untuk mendeskripsikan hubungan antara variabel terikat yang memiliki dua kategori atau lebih dengan satu atau lebih variabel bebas, berskala kategorik atau kontinu. Bentuk umum model LR adalah sebagai berikut (Sartono 2010)

$$x = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}$$

$x$  adalah probabilitas dari LR dan  $\exp$  merupakan fungsi eksponen yang merupakan kebalikan dari fungsi logaritma natural.  $\beta_0$  merupakan konstanta dan  $x_i, i = 1, 2, \dots$  merupakan variabel independen.  $\beta$  merupakan koefisien determinasi pada model LR untuk  $i = 1, 2, 3, \dots$

Salah satu cara untuk mengelompokkan data adalah dengan menggunakan Regresi Logistik Biner. Metode Regresi Logistik Biner merupakan metode klasik yang digunakan untuk mengetahui pola hubungan antara variabel respon yang bersifat biner yakni terdiri dari 0 dan 1 dengan variabel prediktornya (Hosmer *et al* 2000). Hasil dari variabel respon yang terdiri dari 2 kategori yaitu sukses dan gagal yang dinotasikan dengan 1 (sukses) dan 0 (gagal). Untuk kasus peneliti diasumsikan 1 untuk sentimen positif dan 0 untuk sentimen negatif.

### Support Vector Machine

*Support vector machine* (SVM) merupakan salah satu metode dari *supervised learning* yang dapat diaplikasikan untuk kasus klasifikasi, regresi dan deteksi penculan. SVM bekerja dengan melakukan konstruksi *hyperplane* untuk memisahkan data berdasarkan margin di dalam suatu ruang berdimensi  $N$ . Dari  $k$  *hyperplane* yang dibangun, akan dipilih suatu *hyperplane* yang mempunyai nilai margin maksimum yang disebut dengan *optimal hyperplane*.

Persamaan SVM untuk klasifikasi dua kelas adalah sebagai berikut

$$y(x) = w^T \phi(x) + b$$

dimana  $\phi(x)$  menyatakan transformasi dari ruang fitur tetap, variabel  $b$  menyatakan bias, variabel  $w$  merupakan vektor normal dari suatu *hyperplane*, dan titik data baru yang dihasilkan dinyatakan dengan  $y(x)$  (Christopher 2006).

Untuk menemukan *optimal hyperplane* dibutuhkan parameter berupa regularisasi, *kernel*, dan gamma. Parameter  $C$  (regularisasi) berfungsi untuk melakukan kontrol terhadap error yang dihasilkan saat klasifikasi. Parameter gamma mempengaruhi besarnya kelengkungan dari suatu kurva yang berfungsi sebagai *decision boundary*. Parameter ini hanya berlaku jika menggunakan *radial basis function* (RBF) *kernel*. Parameter *kernel* berperan dalam melakukan transformasi dari suatu masalah menggunakan beberapa konsep aljabar linear.

Salah satu keunggulan algoritma ini terdapat di kemampuan belajarnya yang tidak bergantung terhadap dimensi ruang dari fitur tetapi menggunakan margin sebagai alat ukur dari kompleksitas hipotesis (Joachims 1998). Oleh karena itu, suatu data yang berdimensi tinggi dapat dilakukan proses generalisasi menggunakan SVM.

### Skor Validasi

Skor validasi merupakan metode untuk mengatasi ragam yang tinggi dari model yang dihasilkan dengan memisahkan data latih menjadi data validasi. Data validasi akan digunakan untuk menguji model yang dihasilkan sebelum model tersebut digunakan pada data uji. Metode ini akan memisahkan data latih menjadi  $K$  bagian. Untuk setiap  $K$ , akan dilakukan pembuatan model untuk  $K - 1$  bagian lainnya dan  $K$  akan digunakan layaknya data uji untuk menghitung akurasi atau *error* yang dihasilkan. Proses ini akan dilakukan untuk  $k = 1, 2, \dots, K$  dan keseluruhan hasilnya akan digabungkan. Skor validasi untuk error yang dihasilkan adalah sebagai berikut.

$$CV(f) = \frac{1}{N} \sum_{i=1}^N L(y_i, f^{-k(i)}(x_i))$$

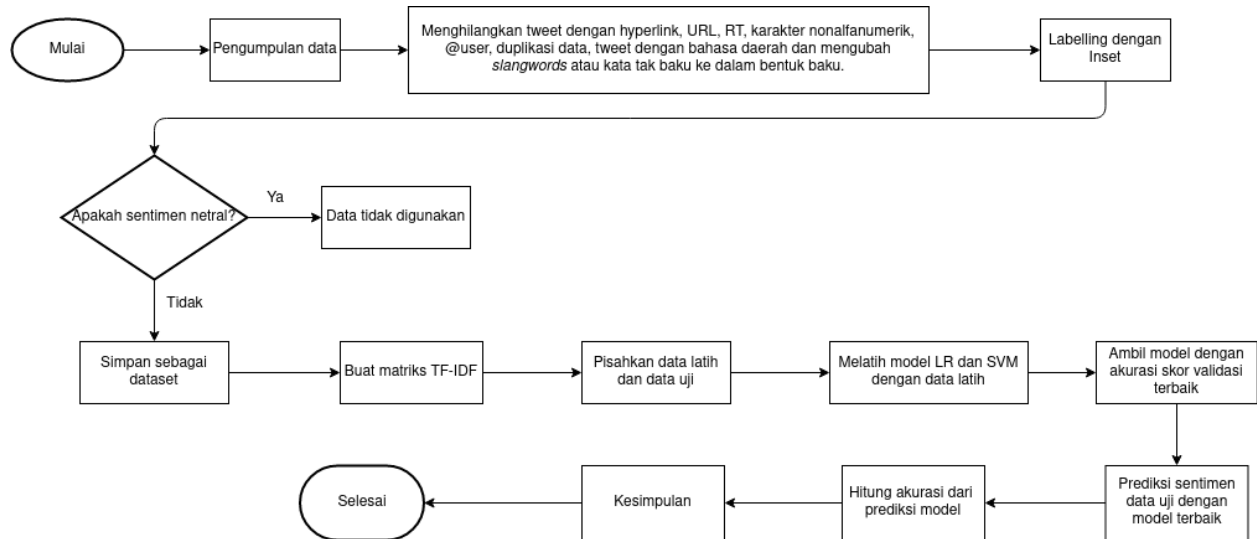
$N$  menyatakan banyaknya pembagian data latih untuk validasi (nilai  $K$ ). Untuk setiap  $i$  sampai dengan  $N$ , dilakukan perhitungan  $L(y_i, f^{-k(i)}(x_i))$  yaitu nilai *error* yang dihasilkan dari setiap prediksi yang dilakukan oleh model dengan menjadikan  $i$  sebagai data validasi. Hasil perhitungan  $L(y_i, f^{-k(i)}(x_i))$  akan digabungkan dengan  $i - 1$  data lainnya dan hasilnya akan dibagi dengan  $N$  untuk mendapatkan rata rata dari skor validasi (Hastie *et al* 2017).  $L(y_i, f^{-k(i)}(x_i))$  dapat pula disubstitusi dengan akurasi dari model yang dihasilkan.

### Inset

InSet merupakan sebuah leksikon sentimen bahasa Indonesia yang memberikan bobot pada setiap kata positif dan negatif dalam rentang nilai -5 hingga 5 (Koto *et al* 2017). Bobot setiap kata akan dijumlahkan untuk menentukan sentimen dari sebuah kalimat. Apabila nilai keseluruhan bobot lebih besar dari 0, maka kalimat tersebut memiliki sentimen positif, begitu pula sebaliknya. Apabila total bobot adalah nol, maka kalimat tersebut adalah kalimat netral.

## METODE ANALISIS DATA

Gambar berikut menunjukkan proses pengerjaan proyek yang dilakukan. Pertama data dikumpulkan melalui *scraping tweet* pada Twitter. Setelah itu dilakukan pembersihan *tweet*. Proses berikutnya adalah labelling sentimen dengan InSet. Data kemudian dipisahkan menjadi data latih dan data uji. Data latih akan dijadikan input untuk model LR dan SVM. Hasil akurasi skor validasi terbaik akan diambil sebagai model untuk memprediksi sentimen data uji dan dihitung akurasinya.

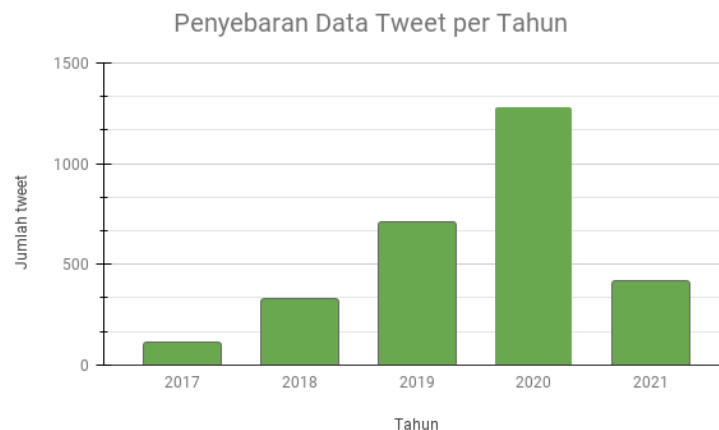


**Gambar 1.** Alur Pengerjaan Proyek

## 1. Pengumpulan Data

Data diambil dari platform twitter dengan berdasarkan hasil pencarian terkait topik peneliti, yaitu alih fungsi lahan pertanian. Peneliti menggunakan *API (Application Programming Interface)* twint untuk mendapatkan data dari Twitter. Data yang sudah didapatkan dari satu *keyword* akan digabungkan dengan data yang diambil dengan *keyword* yang lain. Jumlah data yang didapat adalah 2708.

Bentuk dataset terdiri atas ID pengguna, tanggal, dan *tweet*. Data yang diambil adalah data dalam periode Januari 2017 - Juni 2021. *Keyword* yang peneliti gunakan adalah alih fungsi lahan pertanian, kondisi lahan pertanian, pertanian digusur, lahan pertanian habis, konversi lahan sawah, sawah jadi jalan tol, sawah digusur, keadaan lahan pertanian, kebijakan ketersediaan lahan, penyusutan lahan pertanian, lahan gersang, ladang pertanian, lahan sawah berkurang, konversi lahan pertanian, dan alih fungsi lahan sawah.



**Gambar 2.** Jumlah *tweet* setiap tahun

## 2. Pra Proses Data

Setelah proses pengumpulan data dilakukan, data tersebut harus dilakukan pra proses data untuk menyiapkan data sebaik mungkin sebelum dilakukan *modelling*.

Tahapan awal dari pra proses data yaitu melakukan pembersihan *tweet*. Data tersebut dibersihkan dari link URL, RT, mentions, @user, hashtags, emoji, tag html (&amp, dst), newline, angka dan *punctuations* (simbol). Komponen-komponen tersebut harus dibersihkan dari data *tweet* karena tidak dapat dihitung nilai sentimennya. Adanya komponen tersebut juga membuat proses pelatihan model menjadi lebih sulit karena datanya menjadi sangat berantakan. Karena itu, komponen-komponen tersebut dihapuskan. *Tweet* yang menggunakan bahasa daerah juga dihapuskan dari data karena pembuatan model hanya dilakukan untuk teks bahasa Indonesia.

Setelah itu dilakukan *lowering text* dan tokenisasi untuk menghapus *whitespace* yang tersisa. Berikutnya kata-kata tidak formal (*slang words*) di dalam kolom *tweet* dikumpulkan dan diubah dalam bentuk formalnya menggunakan suatu algoritma penggantian kata yang dibuat peneliti. Duplikasi data yang mungkin muncul setelah dilakukan pembersihan *tweet* dapat dihindari dengan membuang data yang terduplikasi. Hasil dari pembersihan *tweet* secara keseluruhan dapat dilihat pada kolom *clean\_tweet* sebagai berikut.

	tweet	clean_tweet
0	di gua udh ga ada sawah digusur sentraland sem...	di saya sudah tidak ada sawah digusur sentrala...
1	teriak <sup>2</sup> harga properti mahal akhirnya sawah di...	teriak harga properti mahal akhirnya sawah dig...
2	gunungnya udah dikeruk sawah udah jadi pabrik ...	gunungnya sudah dikeruk sawah sudah jadi pabri...
3	semua sawah akan digusur dinggo mbangun dalam ...	semua sawah akan digusur dinggo membangun dala...
4	terpikir jg klo fenomena skrg proyek bangunan ...	terpikir juga kalau fenomena sekarang proyek b...
...	...	...
2869	nggak punya lahan untuk tanaman makanya rumah ...	tidak punya lahan untuk tanaman makanya rumah ...
2870	memberi latihan dan memberi pengalaman secara ...	memberi latihan dan memberi pengalaman secara ...
2871	syoknya tggk pakaian cikguâ <sup>2</sup> ni hariâ <sup>2</sup> boleh pa...	syoknya tegak pakaian cikgu ini hari boleh pak...
2872	fungsi tanggul ini mampu melindungi kawasan la...	fungsi tanggul ini mampu melindungi kawasan la...
2873	tingkatkan proses dimana desa yg membutuhkan a...	tingkatkan proses dimana desa yang membutuhkan...

Gambar 3. Contoh *tweet* yang sudah dibersihkan

Tahap selanjutnya yaitu proses *data labelling* yang bertujuan untuk mengklasifikasikan data berdasarkan sentimennya (positif, negatif, dan netral). Klasifikasi data dilakukan dengan menggunakan InSet.

*Tweet* yang diklasifikasikan sebagai netral oleh InSet tidak akan digunakan dalam tahap *modelling*. Penelitian ini dilakukan untuk membuat model klasifikasi yang dapat memprediksi apakah pendapat seseorang dalam Twitter berkaitan dengan alih fungsi lahan pertanian merupakan pendapat dengan sentimen positif atau negatif, sehingga data netral tidak dimasukkan dalam penelitian.

### 3. Modelling

Tahapan selanjutnya yaitu menjadikan *tweet* yang sudah dibersihkan menjadi matriks *term frequency*. Proses pembuatan matriks *term frequency* menggunakan fungsi TF-IDF yang ada pada pustaka sklearn dan dapat diimplementasikan dengan Python. Hasilnya adalah sebuah matriks *sparse* dengan ukuran 2708×9315.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.200874
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000

**Gambar 4.** Matriks *sparse* hasil fungsi TF-IDF sklearn

Matriks *sparse* tersebut akan dipisahkan menjadi data latih dan data uji. Matriks *sparse* yang digunakan untuk data latih akan dimasukkan sebagai input ke dalam model LR dan SVM. Untuk meningkatkan akurasi dari model, dilakukan *hyperparameter tuning* untuk menentukan parameter terbaik dari model. *Hyperparameter tuning* dilakukan dengan GridSearchCV dari pustaka sklearn dan pustaka Optuna. Parameter terbaik tiap model adalah sebagai berikut.

Model	Parameter
Logistic Regression (LR)	C = 1.113051365843949
Support Vector Machine (SVM)	C = 1.2465148898233047, kernel = linear

**Tabel 1.** Parameter Optimal dari masing-masing model

Parameter tersebut dimasukkan ke dalam model dan dilakukan *cross validation* sebanyak 10 *fold*. Hasil rata rata akurasi skor validasi untuk LR dan SVM ditunjukkan pada tabel berikut.

Model	Skor Validasi
Logistic Regression (LR)	0.80 ± 0.02
Support Vector Machine (SVM)	0.81 ± 0.02

**Tabel 2.** Akurasi skor validasi setiap model



Dari skor validasi data, diketahui bahwa akurasi skor validasi tertinggi dihasilkan oleh model SVM. Model ini akan digunakan dalam tahap *model usage*.

#### 4. Model Usage

Model SVM akan digunakan untuk memprediksi data baru, yaitu data uji sebanyak 813 data yang telah berbentuk matriks *sparse*. Hasil prediksi akan dibandingkan dengan sentimen yang dihasilkan untuk data uji menggunakan InSet. Hasil prediksi model SVM tersebut dapat dilihat menggunakan *confusion matrix*. *Confusion matrix* yang dihasilkan adalah sebagai berikut

	Prediksi Positif	Prediksi Negatif
Sentimen Positif	397 ( <i>True Positive</i> )	80 ( <i>False Negative</i> )
Sentimen Negatif	64 ( <i>False Positive</i> )	272 ( <i>True Negative</i> )

**Tabel 3.** *Confusion matrix* data uji dengan model SVM

Hasil dari *confusion matrix* digunakan untuk menghitung akurasi.

$$\begin{aligned}
 Accuracy &= \frac{TP + TN}{TP + TN + FP + FN} \\
 &= \frac{397 + 272}{397 + 272 + 64 + 80} \\
 &= 0.822
 \end{aligned}$$

Akurasi yang didapat menggunakan model SVM untuk prediksi sentimen alih fungsi lahan pertanian adalah 82.2%.

#### 5. Analisis Hasil

Berdasarkan perbandingan akurasi tiap model pada Tabel 2, SVM mempunyai akurasi skor validasi tertinggi sebesar  $0.81 \pm 0.02$  jika dibandingkan dengan model LR yang memiliki akurasi  $0.80 \pm 0.02$ . Faktor utama yang mempengaruhi tingginya skor validasi model SVM adalah *kernel* yang digunakan yaitu *kernel* linear. *Kernel* linear cocok untuk data teks yang mempunyai banyak fitur sehingga pemetaan data ke ruang dimensi tinggi tidak dibutuhkan. Dengan kata lain, tidak terjadi peningkatan akurasi untuk pemetaan secara *nonlinear* (Chih-Wei *et al* 2003). Data teks memiliki sangat sedikit fitur yang tidak relevan karena data tersebut sudah dilakukan tahap pra proses. SVM juga didukung dengan kemampuan untuk memasang sejumlah *kernel* sehingga fitur tidak relevan dapat dipisahkan dengan jarak yang maksimal. Data teks umumnya dapat dipisahkan secara linear sehingga model SVM cocok digunakan karena model tersebut selalu mencari pemisah linear dari data (Joachims 1998). Faktor lain yang mempengaruhi yaitu parameter C yang merupakan parameter regularisasi. Jika nilai parameter C semakin besar, optimasi akan memilih *hyperplane* yang mempunyai nilai margin yang jauh lebih kecil dan begitu juga sebaliknya.

## KESIMPULAN

Berdasarkan pengujian pada model LR dan SVM, didapatkan bahwa model terbaik untuk memprediksi analisis sentimen alih fungsi lahan pertanian adalah model SVM dengan parameter C bernilai 1.2465148898233047 dan *kernel* yang digunakan adalah *kernel* linear.

## DAFTAR PUSTAKA

- Chih-Wei H, Chih-Chung C, Chih-Jen L. 2003. A Practical Guide to Support Vector Classification. Taipei (TW) : National Taiwan University.
- Christoper M. Bishop. 2006. Pattern Recognition and Machine Learning. Cambridge (UK) : Springer.
- Dunnmon J, Ganguli S, Hau D, Husic B. 2017. Predicting State-Level Agricultural Sentiment with Tweets from Farming Communities.
- Hastie T, Tibshirani R, Friedman J. 2017. The Elements of Statistical Learning: Data Mining, Inference, and Prediction Second Edition. Stanford (US): Springer.
- Hosmer, David W, Lemeshow S. 2000. Applied Logistic Regression Second Edition. New York (US): John Wiley & Sons, Inc.
- Joachims T. 1998. Text Categorization with Support Vector Machines: Learning with Many Relevant Features. European Conference on Machine Learning. 137-142.
- Nandini RA, Sari YA, Adikara PP. 2019. Analisis Sentimen Impor Beras 2018 Pada Twitter Menggunakan Metode *Support Vector Machine* dan Pembobotan Jumlah *Retweet*. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer. 3(4): 3396 - 3406.
- Kompas. 2019. Mentan: Alih Fungsi dari Lahan Pertanian di 2019 Capai 150.000 Hektar. [internet].[diunduh 2021 Juni 30].  
<https://money.kompas.com/read/2021/03/29/140755726/mentan-alih-fungsi-dari-lahan-pertanian-di-2019-capai-150000-hektar>
- Koto F, Rahmaningtyas G. 2017. InSet Lexicon: Evaluation of a Word List for Indonesian Sentiment Analysis in Microblogs. International Conference on Asian Language Processing (IALP). 10.1109/IALP.2017.8300625.
- Republik Indonesia. 2009. Undang Undang No.41 Tahun 2009, tentang Perlindungan Lahan Pertanian Pangan Berkelanjutan.
- Sartono B. 2010. Menduga dan Menguji Koefisien Regresi Logistik Biner Menggunakan Solver di MS Excel.
- Statista. 2021. Leading countries based on number of Twitter users as of April 2021 (in millions). [internet].[diunduh 2021 Juni 28].  
<https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>