Perbandingan Algoritma Cnn Dan Svm Untuk Analisis Sentimen Mengenai Kenaikan Harga Bahan Bakar Minyak

1st Rafly Ahmad Y
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
raflyahmadyanuar@student.telkomuniversity.ac.id

2nd Kemas Muslim L
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
kemasmuslim@telkomuniversity.ac.id

***Abstrak* — Opini-opini maupun keluhan masyarakat yang disampaikan melalui tweet dapat diolah untuk mengetahui sentimen yang ada di dalam tweet tersebut. Pada penelitian ini dilakukan analisis sentimen menggunakan machine learning. Penggunaan machine learning ini dapat mempermudah saat pengambilan data dan pemrosesan data, yang tidak memerlukan banyak waktu dan biaya. Proses klasifikasi data tweet yang dilakukan dalam penelitian ini yaitu data yang mengandung sentimen positif dan sentimen negatif mengenai kebijakan pemerintah yaitu kenaikan harga bahan bakar minyak (BBM). Metode klasifikasi yang digunakan untuk penelitian ini menggunakan Convolutional Neural Network (CNN) dan Support Vector Machine (SVM). Untuk pengambilan data tweet menggunakan metode crawling. Hasil yang didapatkan dari penelitian dengan melakukan evaluasi menggunakan Confusion Matrix mendapatkan bahwa algoritma SVM mendapatkan nilai akurasi yang cukup tinggi sebesar 85% dengan menggunakan max features 510 dan rasio 80:20 dibandingkan dengan algoritma CNN yang memiliki nilai akurasi tertingginya di angka 74% menggunakan nilai max features 300 dan rasio 80:20. Untuk nilai penggunaan cross fold validation CNN mendapatkan nilai rata-rata akurasi tertingginya 78% dengan k=10 sedangkan SVM 87%**

***Kata Kunci:* analisis sentimen, pembelajaran mesin, CNN, SVM, twitter, sosial media**

I. PENDAHULUAN

Bahan Bakar Minyak memiliki peran yang besar bagi masyarakat Indonesia, baik konsumsi secara langsung maupun tidak, karena dampak dari perubahan harga BBM mempengaruhi harga transportasi, produksi, dan lain-lain. Harga bahan makanan pokok juga terpengaruh dari dampak kenaikan harga BBM, contohnya gula, beras, minyak. Kenaikan harga BBM sangat berpengaruh besar bagi kehidupan masyarakat sehingga menimbulkan fenomena pro dan kontra di kalangan masyarakat. Terjadinya kenaikan harga BBM banyak masyarakat yang dirugikan, karena dengan adanya kenaikan harga BBM ini banyak harga bahan pangan dan transportasi juga yang ikut naik. Maka dari itu diperlukan penelitian mengenai opini masyarakat terhadap peristiwa ini, salah satu upaya yang dikembangkan adalah analisis sentimen. Analisis sentimen adalah proses yang digunakan untuk menentukan opini, emosi dan sikap yang dicerminkan melalui teks, dan biasanya diklasifikasikan menjadi opini negatif dan positif[1]. Penelitian analisis sentimen ini dilakukan dengan menggunakan machine learning yang dapat mempermudah penulis untuk pengambilan data, pemrosesan data dan lain-lain. Machine learning juga tidak memerlukan banyak waktu dan biaya, lain halnya dengan melakukan analisis sentimen secara manual pasti akan memerlukan waktu yang lama dan memakan biaya. Analisis sentimen juga merupakan salah satu penelitian yang sering dilakukan sudah banyak penulis yang melakukan penelitian mengenai analisis sentimen, contohnya pada tahun 2020 Prasetiarini T.A., telah melakukan sebuah penelitian analisis sentimen mengenai Maskapai Penerbangan dengan menggunakan metode support vector machine[2]. Pada penelitian kali ini penulis akan memanfaatkan media sosial, yaitu Twitter. Twitter merupakan aplikasi media sosial yang memungkinkan pengguna untuk berinteraksi dengan pengguna yang lain. Twitter juga memberikan akses kepada pengguna untuk mengirimkan pesan singkat yang disebut dengan tweet. Maka dari itu penelitian ini akan mengambil tweet sebagai datanya, karena biasanya masyarakat menggunakan tweet sebagai tempat memberikan komentar tentang peristiwa-peristiwa yang sedang terjadi. Penelitian ini akan menggunakan dua metode klasifikasi, yaitu Convolutional Neural Network dan Support Vector Machine. Pada penelitian ini akan dihasilkan akurasi dari setiap metode, manakah metode yang akan mendapatkan akurasi yang lebih bagus.

II. KAJIAN TEORI

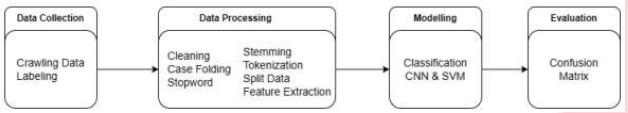
Pada tahap ini, sangatlah penting bagi penulis sebagai referensi untuk mencari tahu keterkaitan antara penelitian yang sudah pernah dilakukan dengan penelitian yang akan dibuat, agar terhindar dari tindakan duplikasi oleh penulis, penelitian terkait yang sudah penulis pelajari dapat disimpulkan pada Tabel 1.

TABEL 1.
Penelitian terdahulu

No	Penulis	Tahun	Metode	Hasil
1	Prasetiarini T.A.	2020	SVM	Kernel Linear: nilai akurasi sebesar 88,75% kernel polynomial: akurasi sebesar 75,625%
2	Saragih P.S	2021	SVM	Nilai akurasi yang diperoleh 85.185%
3	Sartini	2020	CNN	Hasil penelitian performa terbaik pada CNN dengan tingkat akurasi 81,4%
4	Wibowo B.A	2020	CNN	Pada penelitian ini penulis mendapatkan nilai akurasi CNN sebesar 92,62%

III. METODE

Penelitian ini merupakan penelitian kuantitatif. Metode penelitian kuantitatif merupakan suatu cara yang digunakan untuk menjawab masalah penelitian yang berkaitan dengan data berupa angka dan program statistik[3]. Dalam penelitian ini banyak tahapan yang dilakukan, berikut ini tahapan-tahapannya.



GAMBAR 1.
Tahapan penelitian

A. Pengumpulan Data

Selama tahap pengumpulan data, data tweet dikumpulkan dari Twitter menggunakan metode Crawling. Crawling adalah metode pengambilan data dan pengumpulannya untuk analisis. Data yang dikumpulkan terdiri dari tweet-tweet dari Twitter yang terkait kenaikan harga bahan bakar. Proses crawling data dalam penelitian ini dilakukan dengan menggunakan bantuan node.js dan cmd, pada tahapan ini juga akan dilakukan pelabelan data dengan cara manual terhadap data yang telah di dapat.

B. Pemrosesan Data

Pada tahap ini data akan dibersihkan sebelum diproses yang melewati beberapa tahap pembersihan data, seperti di bawah ini:

1. Cleaning: hashtag, mention, dan tanda baca pada proses ini akan dibersihkan.
2. Case Folding: pada proses ini huruf-huruf akan diubah menjadi huruf kecil.
3. Stopword: proses menghilangkan kata-kata yang tidak memiliki makna[4].
4. Stemming: proses untuk menghilangkan imbuhan pada suatu kata.
5. Tokenization: proses yang bertujuan untuk memecah kalimat menjadi kata-kata.

Setelah data dibersihkan, data akan melewati 2 tahap pemrosesan data lain, yaitu Split Data dan Feature Extraction.

C. Modelling

Pada tahap ini dilakukan klasifikasi dengan algoritma support vector machine dan algoritma convolutional neural network. Kedua algoritma ini digunakan karena pada penelitian ini cocok digunakan pada data yang banyak.

D. Evaluasi

TABEL 2.
Confusion matrix

Nilai Prediksi	Kelas Prediksi	
	Positif	Negatif
Positif	TP (True Positif)	FN (False Negatif)
Negatif	FP (false Positif)	TN (True Negatif)

Tahap evaluasi merupakan tahap terakhir yang didalamnya melakukan klasifikasi dengan benar atau tidak, performa text evaluation dapat dievaluasi menggunakan tabel confusion matrix, tabel ini memiliki empat hasil klasifikasi seperti Tabel 2. Confusion matrix ini akan menghitung nilai Accuracy(1), Precision(2), Recall(3), dan F1-Score(4) yang dimana berfungsi untuk mengukur performansi dari model yang digunakan. dibawah ini adalah cara menghitung keempat nilai diatas:

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1-Score = \frac{2 * Recall * Precision}{Recall+Precision} \quad (4)$$

IV. HASIL DAN PEMBAHASAN

A. Pengumpulan data

1. Crawling

Pada pengambilan data ini atau bisa disebut juga dengan crawling data, penulis tidak memakai bantuan library dan API Twitter untuk mengambil data tweet yang diperlukan. Pada penelitian ini penulis menggunakan bantuan dari node.js dan cmd untuk mendapatkan data tweet mengenai kenaikan harga BBM dengan kata kunci “kenaikan bbm’. Berikut adalah visualisasi dari proses pengambilan data tweet dengan bantuan node.js dan cmd:

GAMBAR 2.
Visualisasi crawling data

Keterangan:

a. Command Prompt (CMD)

Sebelum melakukan proses pastikan pc atau laptop anda sudah terinstall node.js, pada proses awal akan dilakukan dalam CMD. mulai dengan penulis melakukan perintah ”npx tweet-harvest@0.0.35”, kemudian akan keluar perintah untuk menginput “auth token” hal ini bisa didapatkan di akun Twitter masing-masing, setelah menginput “auth token” akan ada lagi perintah untuk menginput kata kunci data tweet yang akan diambil, pada penelitian ini penulis memilih kata kuncinya yaitu “kenaikan bbm”. setelah itu akan diperintahkan untuk menginput berapa jumlah data yang akan diambil, penulis membutuhkan 1500 data tweet untuk penelitian ini.

b. Twitter

Pada proses ini sistem akan masuk ke halaman web Twitter akun pengguna dan melakukan sendiri pengumpulan

data tweet dengan cara scrolling halaman Twitter yang berisi kata kunci yang penulis input kan ke dalam cmd, yaitu "kenaikan bbm" dan dengan jumlah data yang diambil 1500 tweet.

c. Excel

Proses terakhir ini sistem akan mengumpulkan data yang telah terambil ke dalam excel dengan format ".csv" dan lalu akan tersimpan sendiri ke dalam folder pengguna.

TABEL 3.
Hasil pelabelan data.

@Herman43955019@tatakujiyati Tolol. Harga BBM naik. Sementara bansos dikaitkan dengan penghasilan dibwh 3.5 jt Lo kira di jateng UMRnya brp? Justru gw pintar. Itu artinya pekerja jateng akan merasakab beratnya kenaikan BBM
jpnncom, "#BeritaJabar Pemkab Cianjur Salurkan BLT Dampak Kenaikan Harga BBM
@abu_waras Udah hancur harga harga di pasar karena kenaikan BBM kemarin apa bisa normal lagi dgn diturunkan lagi?? Mangkanya pake otak dan hati nurani kalau mau bikin keputusan tuh

Tabel 3 adalah contoh hasil dari pengambilan data tweet yang mempunyai kata kunci "kenaikan bbm".

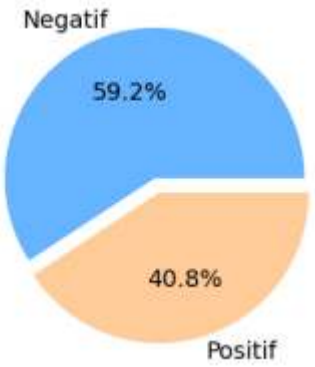
d. Labelling

Pada tahap ini data yang sudah dikumpulkan akan diberi label sentimen, yaitu label Positif dan Negatif. Pelabelan data ini dilakukan secara manual dengan cara melakukan voting, dibutuhkan 3 orang atau berjumlah ganjil untuk melakukan cara voting ini. Setiap orang akan diberi 500 tweet berbeda, karena disini penulis mempunyai data 1500 dan 3 orang untuk melakukan votingnya. Cara melakukannya setiap orang akan membacakan tweet-tweet yang didapat, lalu akan memberikan pendapat masing-masing tentang tweet tersebut apakah bersifat negatif atau positif. Sentimen yang mendapatkan voting terbanyak itulah yang akan dipakai sebagai label untuk tweet itu. Hasil jumlah label bersentimen negatif dan positif dapat dilihat pada Tabel 3.

TABEL 4.
Hasil pelabelan data

Label	Jumlah Data
Positif	550
Negatif	798

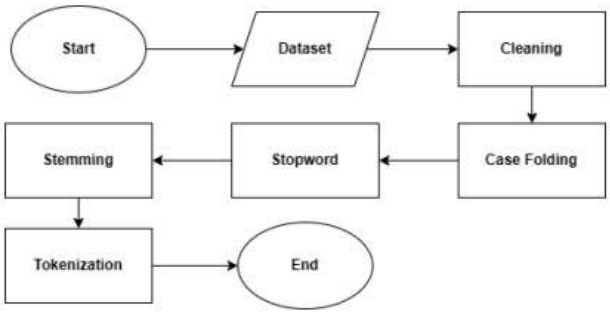
Dilihat dari Tabel 4 data dengan sentimen negatif memperoleh lebih banyak dari pada data dengan sentimen positif, yang bisa divisualisasikan dengan diagram.(Gambar 3).



GAMBAR 3.
Proporsi perbandingan sentiment tweet

B. Pemrosesan Data

Pada pemrosesan data, data yang telah diperoleh akan diproses dulu untuk dibersihkan dari noise. Tahap pemrosesan data memiliki beberapa tahap yang dilakukan, berikut ini tahapan dalam pemrosesan data untuk mendapatkan data bersih:



GAMBAR 4.
Tahapan pemrosesan data

Setelah melalui pembersihan data seperti Gambar 4, akan mendapatkan hasil data bersih dan sudah memiliki label sentimen.(Tabel 5).

TABEL 5.
Hasil pembersihan data

Sebelum	Sesudah	Label Sentimen
@Herman43955019@tatakujiyati Tolol. Harga BBM naik. Sementara bansos dikaitkan dengan penghasilan dibwh 3.5 jt Lo kira di jateng UMRnya brp? Justru gw pintar. Itu artinya pekerja jateng akan merasakab beratnya kenaikan BBM	tolol harga bbm naik sementara bansos kait dengan penghasilan bawah 3 5 kamu kira jateng umr nya berapa justru saya pintar itu artinya pekerja jateng akan merasakan beratnya naik bbm	Negatif
jpnncom, "#BeritaJabar Pemkab Cianjur Salurkan BLT Dampak Kenaikan Harga BBM	pemkab cianjur salurkan blt dampak naik harga bbm	Positif
@abu_waras Udah hancur harga harga di pasar karena kenaikan BBM kemarin apa bisa normal lagi dgn diturunkan lagi?? Mangkanya pake otak dan hati nurani kalau mau bikin keputusan tuh	udah hancur harga harga di pasar karena naik bbm kemarin apa bisa normal lagi dengan turun lagi makannya pake otak dan hati nurani kalau mau bikin putusan tuh	Negatif

Data yang telah dibersihkan akan berubah jumlahnya, karena dalam proses pembersihan data terdapat proses penghapusan data yang sama. oleh karena itu jumlah data awal >1.500 data menjadi 1.348 data tweet. Selanjutnya data akan masuk ke dalam pemrosesan data lain, yaitu Split data, dan Feature Extraction yang memiliki tujuan sebagai berikut:

1. Split Data

Pada proses ini data akan dibagi menjadi dua, yaitu menjadi data latih dan data uji. Data ini akan coba diuji sebanyak 3 kali dengan pembagian rasio sebesar 90:10, 80:20 dan 70:30. Data yang digunakan untuk proses ini adalah data dari kolom 'Text_Clean' dan 'Sentiment'.

2. Feature Extraction

Tahap ini melakukan proses pengubahan data yang sudah melewati proses pembersihan data menjadi data yang divektorisasi, penelitian ini dapat menggunakan metode Term-Frequency - inverse Document Frequency (TF-IDF), yang memiliki tujuan untuk menghitung bobot dari setiap kata-kata yang muncul pada suatu dokumen teks. Pada penelitian ini max features yang digunakan, yaitu 510,400, dan 300

C. Modelling

1. Convolutional Neural Network (CNN)

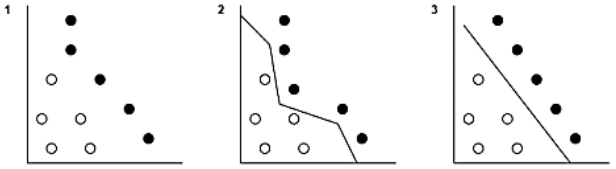
Convolutional Neural Network adalah jaringan syaraf multilayer yang berjenis feed forward network yang memiliki dua atau lebih deep layer dan kemudian dihubungkan dengan fully connected layer seperti multilayer neural network. CNN juga merupakan klasifikasi yang biasanya dilakukan untuk melakukan klasifikasi gambar, mendeteksi objek, pengenalan gambar dan lain-lain. CNN sering dipakai untuk inputan gambar tetapi CNN juga dapat digunakan dalam kalimat atau teks[5].

GAMBAR 6.
Arsitektur CNN

Gambar 6 merupakan visualisasi dari arsitektur CNN. Proses dimulai dengan data yang setiap katanya diubah menjadi vektor lalu di inputkan ke dalam lapisan Convolutional. ConvID digunakan pada lapisan Convolutional ini dengan memakai filter sebanyak 64, dan ukuran untuk kernel yaitu 3. Lalu hasil dari lapisan ini akan diteruskan ke lapisan MaxPooling dengan tujuan untuk mengurangi beban pada komputasi, selanjutnya meneruskan pada lapisan Dropout yang memiliki fungsi untuk menghindari overfitting[6]. Setelah itu diteruskan ke pada lapisan Dense yang memiliki fungsi untuk meneruskan hasil dari Dropout ke pada lapisan Flatten, pada lapisan Flatten dimensi data multiarray akan diubah menjadi vektor. Dan pada lapisan terakhir yaitu lapisan Dense akan melakukan prediksi akhir pada model.

2. Support Vector Machine (SVM)

Support Vector Machine adalah algoritma machine learning yang dapat digunakan pada kasus klasifikasi maupun regresi, SVM merupakan algoritma supervise learning yang memiliki pendekatan yang unik. Klasifikasi yang dilakukan SVM adalah dengan membuat garis pembatas (hyperplane) antara kelas opini positif dan opini negatif untuk memisahkannya. Suatu garis pembatas yang memiliki jarak terbesar ke titik data pelatihan terdekat dari setiap kelas bisa dibilang adalah garis pembatas yang baik, karena pada umumnya semakin besar margin, semakin rendah error generalisasi dari pemilah. Jarak dari suatu titik vektor di suatu kelas terhadap hyperplane disebut juga Margin[7].



GAMBAR 7.
Arsitektur SVM

Gambar 7 adalah visualisasi arsitektur dari algoritma SVM. Gambar yang bernomor 1 adalah ilustrasi dari sebuah data yang didalamnya memiliki 2 kategori, seperti penelitian ini yaitu memiliki kategori positif dan negatif. Kemudian kedua kategori itu dipisahkan oleh sebuah kurva seperti gambar bernomor 2. Dan pada gambar bernomor 3 data telah dilakukan transformasi menggunakan fungsi kernel yang memiliki tujuan untuk mengklasifikasi data non-linear dengan cara mengubahnya menjadi data linear, setelah itu membentuk hyperplane seperti gambar bernomor 3.

D. Evaluasi

Pada bab ini akan melakukan evaluasi hasil dari penelitian dengan menggunakan Confusion Matrix, dimana kali ini penulis melakukan eksperimen dengan algoritma CNN dan SVM. Penulis mempunyai tiga skenario untuk setiap algoritma. Skenario pertama penulis melakukan percobaan dengan membagi data latih dan data uji dengan rasio 90:10, 80:20, dan 70:30. Kedua penulis melakukan percobaan dengan mengubah nilai dari max features TF-IDF dengan nilai 510, 400, dan 300. Ketiga melakukan perhitungan rata-rata akurasi dengan menggunakan cross fold validation, dengan k=10 dan k=5. Hal ini dilakukan untuk mengetahui hasil akurasi dari setiap skenario, apakah akan mendapatkan nilai akurasi berbeda-beda atau mendapatkan nilai akurasi yang sama dan untuk mengetahui juga pada skenario mana algoritma mendapatkan nilai akurasi tertingginya.

Pada skenario pertama mendapatkan hasil dari setiap rasio pada metode CNN dan SVM yang dimana untuk nilai akurasi yang paling tinggi untuk algoritma CNN terdapat pada rasio 80:20 dengan nilai akurasi 78%. Sedangkan untuk algoritma SVM pada rasio 70:30 dan 80:20 mendapatkan nilai akurasi yang sama, yaitu 85%. Untuk selengkapnya hasil semua akurasi akan disatukan pada Tabel 6.

TABEL 6.
Hasil semua rasio

Rasio	CNN	SVM
70:30	73%	85%
80:20	74%	85%
90:10	73%	83%

Skenario kedua melakukan percobaan dengan mengubah nilai max features TF-IDF. Rasio yang digunakan untuk percobaan ini 80:20, karena pada rasio ini kedua algoritma mendapatkan akurasi tertingginya. Hasil yang didapatkan dari setiap nilai berbeda-beda, untuk algoritma CNN nilai akurasi tertingginya jika memakai nilai max features 300 sedangkan untuk SVM pada nilai 510. Hasil semua akurasi dari setiap nilai akan disatukan pada Tabel 7.

TABEL 7.
Hasil semua max features

Nilai	CNN	SVM
510	73%	85%
400	72%	84%
300	74%	83%

Untuk skenario ketiga mengetahui hasil dari penggunaan cross fold validation dengan k=10 dan k=5. Didapatkan hasil dari percobaan ini, yaitu untuk algoritma CNN mendapatkan rata-rata akurasi tertinggi di 78% menggunakan k=10. Sedangkan untuk SVM mendapatkan 87% untuk k=10. Hasil nilai rata-rata akurasi dari setiap algoritma akan disatukan pada Tabel 8.

TABEL 8.
Hasil penggunaan cross fold validation

Nilai k	CNN	SVM
k=10	78%	87%
k=5	77%	86%

V. KESIMPULAN

A. Kesimpulan

Berdasarkan hasil dari penelitian ini dapat ditarik kesimpulan bahwa:

1. Skenario pertama dengan percobaan mengubah nilai dari setiap rasio mendapatkan bahwa, CNN mendapatkan nilai akurasi tertingginya pada rasio 80:20 dengan akurasi 74%, sedangkan SVM pada rasio 80:20 dan 70:30 yang memiliki akurasi sebesar 85%
2. Pada skenario kedua yaitu, melakukan pengubahan nilai max features TF-IDF dengan menggunakan rasio 80:20. Didapatkan nilai akurasi tertinggi CNN pada max features 300 dengan akurasi 74%, dan untuk SVM pada max features 510 dengan nilai akurasi 85%. 3. Untuk skenario ketiga dilakukan percobaan menghitung rata-rata akurasi menggunakan cross fold validation. Mendapatkan hasil untuk CNN dengan k=10 sebesar 78% dan k=5 77%. Sedangkan SVM k=10 mendapatkan 87% dan k=5 86%. 4. Pada penelitian ini algoritma SVM lebih baik dari algoritma CNN, ini mungkin bisa terjadi dikarenakan data yang digunakan terbilang sedikit. Karena SVM memiliki waktu proses yang singkat maka dari itu sangat cocok untuk data yang sedikit.

B. Saran

Hasil penelitian ini menimbulkan saran yang dapat dipakai untuk penelitian selanjutnya:

1. Penelitian selanjutnya dapat menggunakan dataset yang lebih banyak, dengan memakai sentimen berlabel netral tidak hanya positif, negatif nya saja.
2. Dapat menggunakan arsitektur algoritma CNN yang lain, untuk mengetahui hasil dari setiap arsitektur mana yang lebih baik untuk penelitian analisis sentimen

REFERENSI

[1] Coletta, L.F.S.”Combining Classification and Clustering for Tweet Sentiment Analysis”, 2014. Brazilian Conference on Intelligent Systems. IEEE, pp.210-215

[2] Prasetiarini, T.A., Ermawati, I. & Chamidah, N. “Analisis Sentimen Media Sosial Twitter Menggunakan Metode Support Vector Machine(SVM)(Studi Kasus: maskapai Penerbangan PT Garuda Indonesia (PERSERO)TBK)”,2020.

[3] Mp. Dosen Fakultas Ilmu Tarbiyah dan Keguruan UIN Maulana Malik Ibrahim Malang, “PEMAPARAN METODE PENELITIAN KUALITATIF,” 2017.

[4] Luqyana, W.A., Cholissodin, I. & Perdana, R.S. “Analisis Sentimen Cyberbullying pada Komentar Instagram Dengan Metode Klasifikasi Support Vector Machine”,2018.

[5] Britz, D. “Understanding Convolutional Neural Network for NLP”,2015.

[6] R, K, Kaliyar, A. Goswami, P. Narang, and S. Sinha, “FNDNet - A deep convolutional neural network for fake news detection. “Cogn. Syst. Res., vol. 61, pp. 32-44, 2020, doi: 10.1016/j.cogsys.20019.12.005.

[7] Santoso, V.I., Virginia, G. & Lukiko, Y. “Penerapan Sentiment Analysis pada Hasil Evaluasi Dosen Dengan Metode Support Vector Machine”, 2017. Jurnal Transformatika 14(2), 72.

[8] Wibowo, B.A. “Analisis Sentimen Data Twitter Dengan Metode Deep Learning Menggunakan Algoritma Convolutional Neural Network Pada Pilpres 2019”,2020.

[9] Saragih, P.S., Witarsyah, D. & hamami, F. “Analisis Sentimen pada Media Sosial Twitter Menggunakan Algoritma SVM (Studi Kasus: PSBB pada Masa Pandemik Covid-19 Di Provinsi DKI Jakarta”,2021.

[10] Nasukawa, T. & Yi, J. “Sentiment Analysis: Capturing Favorability Using Natural Language Processing”,2003. Proceedings of the 2nd International Conference on Knowledge Capture,pp. 20-77

[11] Nugroho, A.S. “Pengantar Support Vector Machine”,2007. Hitech Research (HRC) from Ministry of. Educatio, Culture, Sports, Science and Technology. Nagoya Inst of Technology, Japan.

[12] Mayer, D. “Support Vector Machine”,2015. Scientific Research An Academic Publisher, e1071.

[13] Yuliska., Hidayatul, D., Lubis, J.H., Syaliman, K.U. & Najwa N.F. “ Analisis Sentimen pada Data Mahasiswa Terhadap kinerja Departemen Di perguruan Tinggi Menggunakan Convolutional Neural Network”, 2021. Jurnal teknologi Informasi dan Ilmu Komputer (JTIIK), Vol. 8, no 5.

[14] Kim, Y. “Convolutional Neural Network for sentence classification”,2014. EMPL 2014 - 2014 Conferences on Empirical Methods in Natural Language Processing, Proceedings of the conference. hal. 1746-1751.

[15] Giummole, F., Orlando, S. & Tolomei. “Trending Topics on Twitter Improve the Prediction of Google Hot Queries”,2013. International Conference on Social Computing. IEEE,pp. 39-44

[16] Sartini. “Analisis Sentimen Twitter Bahasa Indonesia Menggunakan Algoritma Convolutional Neural Network”,2020