

Perbandingan Kinerja CNN dan Naïve Bayes pada Analisis Sentimen Performa Manchester United di Twitter

Kurnia Muludi¹, Muhammad Naufal Humam², Dewi Asiah Shofiana³, Admi Syarif⁴

^{1,2,3,4}Jurusan Ilmu Komputer, Universitas Lampung
²nhumam86@gmail.com

Abstrak— Analisis sentimen digunakan untuk mengevaluasi kecenderungan suatu opini atau sentimen, termasuk apakah pendapat tersebut cenderung memberikan opini yang negative maupun positif. Pada studi penelitian terbaru, data didapatkan dari media social seperti Twitter, di mana pengguna memberikan pendapat pribadi mereka tentang topik tertentu. Manchester United menjadi klub yang sering dibicarakan di twitter mengenai performanya dalam sepakbola, baik dalam performa bagus ataupun performa buruk. Penelitian ini membandingkan dua algoritme, yaitu *Convolutional Neural Network (CNN)* dan *Naive Bayes* untuk menganalisis sentimen data sentiment dari Twitter tentang Manchester United. Data pada penelitian ini berbahasa Inggris dan Indonesia. Berdasarkan hasil penelitian, pada dataset berbahasa Inggris CNN menghasilkan kinerja terbaik dengan nilai akurasi 94%, sementara nilai *Naive Bayes* hanya 79%. Begitupun pada dataset berbahasa Indonesia, CNN menghasilkan kinerja terbaik dengan nilai akurasi 91%, sementara hasil nilai *Naive Bayes* ada pada 75%. Untuk sentimen soal Manchester United, sentimen positif cenderung muncul ketika hasil pertandingan dimenangkan oleh Manchester United, dan pada sentimen berbahasa Inggris maupun Indonesia sentimen positif banyak diisi oleh kata-kata pujian. Sementara sentimen negatif cenderung muncul ketika Manchester United kalah dalam pertandingan atau ketika pemain bermain buruk dan berperilaku indisipliner.

Kata Kunci— Analisis Sentimen, Manchester United, Twitter, CNN, Naive Bayes.

I. PENDAHULUAN

Analisis sentimen merupakan sekumpulan metode, alat, dan pendekatan yang digunakan untuk menemukan dan mengekstrak informasi yang subjektif seperti pendapat dari suatu tulisan [1]. Analisis sentimen juga dapat digunakan untuk mengevaluasi suatu pendapat, termasuk apakah pendapat tersebut cenderung memberikan opini positif atau negatif.

Banyak orang menyukai sepak bola. Pencinta sepak bola tidak melihat usia, ras, suku, agama, golongan, atau jenis kelamin. Banyak dari fans ini mendukung klub-klub dari liga utama Eropa seperti Manchester United, Juventus, dan Barcelona, bukan hanya klub lokal. Fans adalah mereka yang mendukung dari jauh dan biasanya menggunakan media untuk mendukung dan menyaksikan pertandingan tim kesayangannya [2]. Analisis sentimen dapat digunakan dalam olahraga sepak bola untuk menganalisis argumen seseorang yang mendukung suatu klub sepak bola terhadap kinerja klub tersebut.

Manchester United Football Club adalah klub sepak bola profesional Inggris yang berbasis di Old Trafford, Manchester Raya, dan bermain di Liga Inggris. Menurut Deloitte, perusahaan konsultan akuntansi dan industri olahraga, Manchester United memiliki 75 juta pendukung global. Ada juga yang memperkirakan bahwa ada lebih dari 333 juta pendukung [3]. Dengan jumlah penggemar yang sangat

banyak, Manchester United menjadi klub yang sering dibicarakan mengenai performanya dalam permainan sepakbola, baik ketika dalam performa yang bagus ataupun performa yang buruk.

Untuk analisis sentimen, ilmu statistika memiliki banyak metode klasifikasi, tetapi metode Naïve Bayes Classifier (NBC) adalah yang paling umum digunakan untuk klasifikasi teks. Selain menggunakan *naive bayes*, algoritme *deep learning* juga digunakan untuk mengklasifikasikan teks, algoritme *convolutional neural network (CNN)* adalah salah satu algoritme yang dapat digunakan untuk klasifikasi teks. Dalam penelitian analisis sentimen terhadap film melalui media twitter, akurasi tertinggi diperoleh algoritme CNN sebesar 80,99% [4]. Penelitian mengenai pengembangan aplikasi sentiment analysis menggunakan metode Naive Bayes. Studi ini menggunakan *case folding*, *parsing*, dan transformasi pra-proses teks, yang menghasilkan akurasi klasifikasi 73%. [5]. Selain itu, penelitian lain yang berjudul *Indonesian news classification using convolutional neural network* yang juga menggunakan metode CNN menyimpulkan bahwa algoritme CNN dapat digunakan untuk mengolah data teks selain digunakan untuk memproses data citra. CNN dapat melakukan klasifikasi berita Indonesia dengan rata-rata hasil pengujian sekitar 90,74% [6].

Berdasarkan temuan dari beberapa studi sebelumnya, maka penelitian ini bertujuan untuk membandingkan dua algoritme, yaitu CNN, dan juga algoritme yang sudah sangat sering digunakan dalam klasifikasi, *naive bayes* untuk mengetahui kinerja algoritme dengan tingkat efektivitas, efisiensi, dan akurasi paling baik untuk mengklasifikasikan sentimen para pendukung Manchester United.

II. TINJAUAN PUSTAKA

A. Convolutional Neural Network (CNN)

Dengan tujuan mengolah data dua dimensi, *Convolutional Neural Network (CNN)* adalah pengembangan dari *Multilayer Perceptron (MLP)*. Karena kedalaman jaringannya yang tinggi, CNN termasuk dalam kategori *Deep Neural Network* [7]. CNN biasanya digunakan untuk mengklasifikasikan gambar, video, atau dua dimensi. Selain itu CNN dapat digunakan untuk mengklasifikasikan data teks [8]. CNN dapat mengenali informasi prediktif dari teks, potongan suara, dan gambar karena arsitekturnya.

B. Naive Bayes

Algoritme Naive Bayes digunakan untuk menemukan nilai probabilitas tertinggi untuk mengklasifikasikan data uji ke dalam kategori yang paling tepat [9]. Metode klasifikasi *naive bayes* merupakan metode yang dapat digunakan untuk mengklasifikasikan teks. Kelebihan Algoritme *Naive Bayes* adalah algoritmenya, Meskipun sangat sederhana, algoritma ini

sangat akurat. Klasifikasi tweet dilakukan dalam dua tahap. Tweet yang telah diketahui kategorinya dilatih, dan tweet yang belum diketahui kategorinya diklasifikasikan berdasarkan data yang telah dilatih [10]. Metode ini menggunakan pengalaman sebelumnya untuk memprediksi kemungkinan di masa depan. Nilai fitur tertentu tidak bergantung pada nilai fitur lain, menurut klasifikator Naive Bayes, karena variabel kelas [11].

Naive Bayes memiliki tingkat kesalahan paling sedikit dibandingkan dengan metode klasifikasi lainnya. Namun, asumsi yang dibuat untuk penggunaannya, seperti kondisi kelas independen, dan kekurangan data probabilitas yang tersedia, menyebabkan hal ini tidak selalu terjadi dalam praktiknya. Pengklasifikasi bayesian juga bermanfaat karena memberikan pembenaran teoritis untuk pengklasifikasi tambahan yang tidak menggunakan teorema bayes secara eksplisit.

C. Text Mining

Text mining adalah penggalian data yang digunakan untuk menyelesaikan masalah kebutuhan informasi dengan menggunakan teknik penggalian data, pemrosesan bahasa alami, pencarian informasi, dan manajemen pengetahuan. Text mining melibatkan praproses dokumen seperti kategorisasi teks, ekstraksi informasi, dan ekstraksi kata. Metode ini digunakan untuk mengekstraksi data dari sumber data dengan mengidentifikasi dan menyelidiki pola yang menarik [12]. Dalam menangani masalah klasifikasi, *clustering*, ekstraksi data, dan *information retrieval*, *text mining* adalah teknik yang digunakan [13].

Pada dasarnya, penelitian *text mining* mengambil banyak dari proses kerja *data mining*, tetapi pola yang digunakan dalam *text mining* berasal dari kumpulan bahasa alami yang tidak terstruktur, sedangkan pola dalam *data mining* berasal dari *database* yang terstruktur.

D. Analisis Sentimen

Analisis sentimen adalah jenis penelitian yang melihat pendapat, sentimen, evaluasi, penilaian, sikap, dan emosi terhadap sesuatu seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik, dan atributnya. Analisis sentimen sama dengan analisis media sosial. Dewasa ini, analisis sentimen berpusat pada media *online*. Sebab itu, penelitian tentang analisis sentimen mempengaruhi penggunaan NLP serta berdampak besar pada ilmu manajemen, politik, sains, ekonomi, dan ilmu social secara keseluruhan karena pendapat orang memengaruhi semua bidang ini [14].

Salah satu tugas utama dalam analisis sentimen adalah mengklasifikasikan data teks dan kemudian menetapkan apakah opini yang diungkapkan pada kalimat atau dokumen itu positif atau negatif. Analisis sentiment dapat mengungkapkan perasaan marah, gembira, atau sedih. Seseorang dapat menggunakan internet untuk mengumpulkan ulasan tentang

merek, produk, atau orang dan menentukan apakah mereka dianggap positif atau negatif.

E. Twitter

Twitter Inc membangun dan mengelola situs *web* yang menyediakan jaringan social yang disebut Twitter, berupa microblog yang memungkinkan penggunanya untuk membaca dan mengirim pesan *tweet* [15]. Mikroblog merupakan jenis alat untuk berkomunikasi melalui internet dimana seseorang membagikan aktivitas yang sedang dilakukan ataupun pendapat tentang suatu objek dan fenomena, dengan cara menulis status yang disebut *tweet*. *Tweet* adalah teks berupa tulisan hingga 140 karakter yang ditampilkan pada halaman profil pengguna. Namun sejak tahun 2020 jumlah batas maksimal karakter twitter sudah ditambah menjadi 280 karakter. Pengguna twitter dapat membatasi tweet nya untuk bisa dilihat secara publik ataupun secara privat ke daftar orang orang yang telah ditentukan.

F. Manchester United

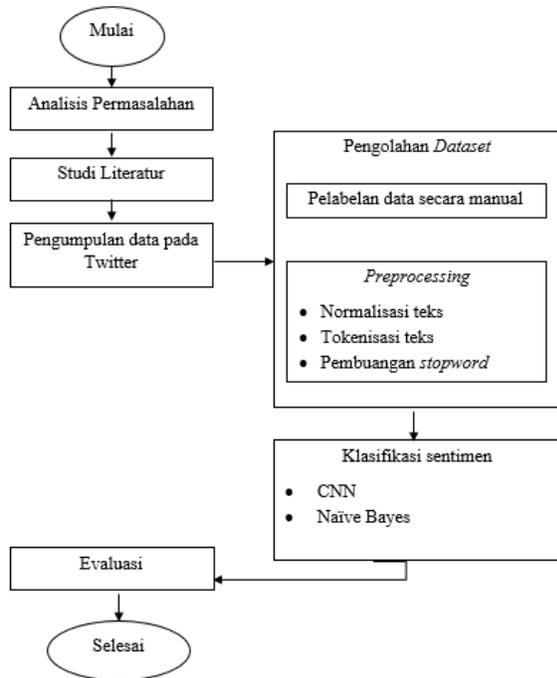
Manchester United (MU) merupakan salah satu tim terbaik di liga primer Inggris, dan Manchester United adalah salah satu tim yang paling sungguh-sungguh dalam mengiklankan dan mengelola media mereka [16]. Di antara klub liga inggris lain Manchester United berada di urutan teratas untuk penggemar paling banyak dengan total lebih dari 91 juta fans di Facebook dan Twitter. Manchester United berada di posisi teratas dengan memiliki sekitar 43 juta fans di Facebook dan Twitter jauh meninggalkan Liverpool yang berada di peringkat kedua.

III. METODOLOGI PENELITIAN

Penelitian ini membandingkan dua algoritme klasifikasi yaitu, Convolutional Neural Network dan Naive Bayes untuk mengetahui bagaimana kinerja kedua algoritme tersebut mengelompokkan tweet yang dibuat oleh pengguna Twitter bernilai positif atau negatif, dinilai pada *accuracy*, *precision*, *recall*, dan *f-1 score*. Kedua algoritme tersebut yang digunakan untuk menganalisis sentimen *tweet* performa Manchester United. Sumber data penelitian ini diambil dari twitter menggunakan *library twint*, sebuah *library* untuk menscraping data *tweet* dari Twitter. Twint menggunakan operator pencarian pada Twitter untuk mengumpulkan data dari user tertentu dan juga dapat mengumpulkan tweet yang berkaitan dengan topik, *hashtag* dan *trending topic*. Twint juga dapat memilah dari informasi yang sensitif dari twitter seperti *e-mail* dan nomer telepon *user*.

A. Tahapan Penelitian

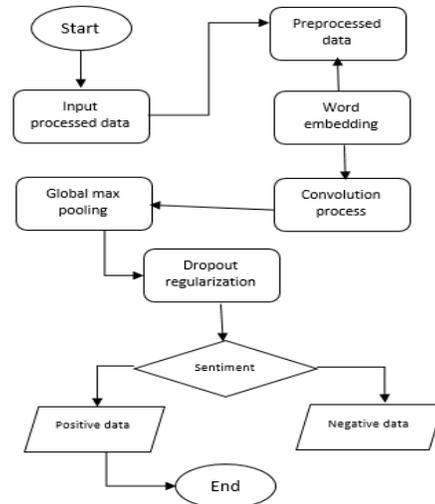
Penelitian ini dilakukan melalui beberapa langkah, seperti yang ditunjukkan pada diagram alir penelitian berikut :



Gbr 1. Diagram Alir Penelitian

Pada Gbr 1 dijelaskan bahwa alur penelitian dimulai dari analisis permasalahan, yaitu membandingkan dua algoritme klasifikasi pada kasus sentimen performa Manchester United yang didapatkan dari twitter. Kemudian dikumpulkan data dari penelitian sebelumnya yang berkaitan dengan subjek penelitian ini. Data yang digunakan dalam penelitian ini adalah data teks dari kumpulan *tweet* dari pengguna twitter soal performa Manchester United di seluruh kompetisi sepakbola baik domestik maupun Eropa seperti *Champions League* dan *Europa League*. Setelah data terkumpul, data kemudian masuk kedalam tahap praproses. Tujuannya untuk mengubah data yang telah didapatkan menjadi bentuk data yang sesuai sebelum masuk ke tahap implementasi.

Selanjutnya pada tahap klasifikasi adalah implementasi CNN dan *Naive Bayes*. Data yang sebelumnya telah melalui praproses akan dikategorikan menjadi sentimen positif atau negatif menggunakan algoritme CNN dan *Naive Bayes*. Kemudian data yang sudah melalui tahap praproses, akan diolah menggunakan algoritme CNN yang diantaranya terdapat proses konvolusi dan *max-pooling*. Pada tipe data dua dimensi atau gambar, operasi ini akan diulang sebanyak jumlah *pixel* input gambar. Karena jenis data yang digunakan dalam penelitian ini adalah teks atau data satu dimensi, teks harus diubah terlebih dahulu menjadi vektor melalui proses *word embedding*. Setelah proses tersebut data diklasifikasikan ke dalam dua kelas yaitu kelas positif, dan kelas negatif. *Flowchart* untuk implementasi algoritme CNN dapat dilihat pada gambar berikut :



Gbr 2. Flowchart Implementasi Algoritme CNN

Kemudian untuk algoritme *Naive Bayes*, setelah data melalui praproses, berikutnya data akan dilatih untuk mengumpulkan informasi dalam bentuk nilai probabilitas kata pada suatu sentimen baik negatif maupun positif. *Naive Bayes* memiliki persamaan berikut ini :

$$P(H | X) = \frac{P(X|H)P(H)}{P(X)} \quad (1)$$

X = Data dengan kelas yang belum diketahui
 H = Hipotesis data X adalah kelas khusus
 $P(H|X)$ = Probabilitas hipotesis H berdasarkan kondisi x
 $P(H)$ = Probabilitas hipotesis H
 $P(X|H)$ = Probabilitas X berdasarkan kondisi tersebut
 $P(X)$ = Probabilitas dari X

Kemudian akan dilakukan proses klasifikasi menggunakan Algoritme *Naive Bayes* terhadap *data testing* dengan *data training* tadi sebagai acuan pengetahuannya. Lalu data juga akan dibagi ke beberapa kelas antara lain kelas positif, dan kelas negatif.

IV. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Dalam proses penelitian ini, *dataset* dikumpulkan dari media sosial Twitter. Data *tweet* yang digunakan adalah *tweet* yang mengandung kata kunci yang berhubungan dengan Manchester United, seperti nama para pemain dan hasil pertandingan yang telah dimainkan. Bahasa yang digunakan pada data untuk penelitian ini berupa Bahasa Indonesia dan Bahasa Inggris. *Dataset* yang dibutuhkan adalah teks *tweet* berisi komentar soal Manchester United yang kemudian dimuat ke *file comma separated value (csv)*. Pelabelan teks *tweet* menjadi positif dan

negatif pada dataset ini dilakukan secara manual dengan menilai isi

komentar pada teks tweet tersebut, kemudian ditentukan apakah komentar tersebut bernilai positif atau negatif. Tweet yang bernilai positif maupun negatif dipengaruhi oleh performa Manchester United. Ketika hasil pertandingan dimenangkan oleh Manchester United, dan para pemain bermain bagus pada pertandingan itu, maka tweet cenderung bernilai positif. Sedangkan ketika Manchester United kalah pada suatu pertandingan, ataupun dari pemain berperilaku tidak baik seperti tindakan indisipliner, tweet yang didapatkan cenderung bernilai negatif. Dataset yang digunakan berjumlah 2.018 data teks dengan data berbahasa Indonesia berjumlah 1.016 dan data berbahasa Inggris berjumlah 1.002. Contoh dari data yang sudah dikumpulkan bisa dilihat pada TABEL I dan TABEL II.

TABEL I
 CONTOH DATASET BERBAHASA INGGRIS

No	Tweet	Sentimen
1	It's great to see United go on to the next round getting a cup win this season would be fantastic. And Garnacho man had a great game today! #mufc	positif
2	What a wonderful game, I enjoyed that one. Big time!! #mufc #MUNAVL	positif
3	What amazing match Garnacho McTominay Kroos Rashford Martial Sancho Maguire Dalot ðŸ˜¸ Villa #FanDuelisELITE #Santaboy #WakandaForever #MUFC Disney quiz https://t.co/gn71QOa58O	positif
4	First domestic cup tie and first win for the boss â€¦ #MUFC #CarabaoCup	positif
5	Dad and son night tonight and a trip to Old Trafford. #MUFC #DadAndLad https://t.co/XrT3zxkvHj	positif
6	Our players do not know how to press ðŸ˜¸ #mufc	negatif
7	Well that escalated quickly #MUFC	positif

TABEL II
 CONTOH DATASET BERBAHASA INDONESIA

No	Tweet	Sentimen
1	@idextratime Ya terus apa min? Dia kan begitu gara2 no respect from the club jadi wajarlah tindakannya begitu. Inget min Ronaldo>	Negatif
2	kemenangan emyu ketutup sm podcast ronaldo uu	Negatif
3	gw suka emyu gara2 rooney dari 2012 mana pas bgt itu season terakhir emyu juara epl bangke kkwkwkwkw	Positif

4	@PorosHalangID Aku suka Bang do sebagai pemain. Dan aku suka Emyu sebagai Klub. Awal nya seneng duo favorit jdi 1 Tpi keknya skrng lebih baik Di pisah aja deh ðŸ˜¸ ðŸ˜¸	Negatif
5	@GOAL_ID Kesalahan terbesar saat menjelang revolusi emyu adalah membeli ronaldo.	Negatif
6	@PorosHalangID Jadi fans emyu pengen seneng dikit gaboleh kayanyaa:)	Negatif
7	Ronaldo udah ga dianggap legend sama fans Emyu gara2 blak2an jelek2in tim di interview-nya Piers Morgan. Gapapa...ente kan emang legend-nya Madrid. Pencapaian terbesar karier kan waktu main di Spanyol ðŸ˜¸	Negatif

B. Pembagian Data

Data yang sudah melalui proses pelabelan secara manual selanjutnya akan dipisahkan menjadi *data training* dan *data testing*. Pembagian data latih sebesar 80% dan untuk data uji sebesar 20%. Pembagian data ini dilakukan secara acak oleh sistem klasifikasi. Ketika ukuran data besar, algoritme pembagian data tidak mempengaruhi hasil. Namun, ketika menggunakan data kecil, algoritme pembagian data dapat meningkatkan performa generalisasi. Hal ini menyimpulkan bahwa tidak ada bukti yang pasti yang menunjukkan kombinasi metode dan parameter mana yang selalu memberikan hasil yang signifikan lebih baik dari yang lain. Namun, keseimbangan yang baik antara ukuran set pelatihan, validasi, dan pengujian dapat memberikan estimasi yang stabil terhadap performa model [17]. Pembagian untuk data latih dan data uji dapat dilihat pada TABEL III dan Tabel IV.

TABEL III
 PEMBAGIAN DATA BERBAHASA INGGRIS

Kelas	Data Latih	Data uji
Positif	510	128
Negatif	291	73
Total	801	201

TABEL IV
 PEMBAGIAN DATA BERBAHASA INDONESIA

Kelas	Data Latih	Data uji
Positif	467	117
Negatif	345	87

Total	812	204
-------	-----	-----

C. *Praproses Data*

Sebelum mengklasifikasikan data yang sudah dilakukan proses pelabelan, data harus terlebih dahulu melalui tahap preprocessing. Tahap yang pertama dilakukan pada praproses data adalah menormalisasikan teks dengan melakukan melakukan pembuangan tanda baca, mengubah huruf-huruf kapital menjadi huruf kecil, dan membuang karakter-karakter khusus maupun simbol-simbol, agar nantinya tidak terjadi kesalahan dalam penafsiran teks. Contoh hasil normalisasi teks pada data berbahasa Inggris dan Indonesia bisa dilihat pada tabel V.

TABEL V
 HASIL NORMALISASI TEKS BERBAHASA INGGRIS DAN INDONESIA

Train tweet	Norm train tweet
It's great to see United go on to the next round getting a cup win this season would be fantastic. And Garnacho man had a great game today! #mufc	it is great to see united go on to the next round getting a cup win this season would be fantastic and garnacho man had a great game today mufc
What a wonderful game, I enjoyed that one. Big time!! #mufc #MUNAVL	what a wonderful game i enjoyed that one big time mufc munavl
Kesalahan terbesar saat menjelang revolusi emyu adalah membeli Ronaldo.	kesalahan terbesar saat menjelang revolusi emyu adalah membeli ronaldo
Berita kemenangan emyu lenyap karena berita Ronaldo.	berita kemenangan emyu lenyap karena berita ronaldo

Setelah melalui proses normalisasi, selanjutnya dilakukan tokenisasi dan pembuangan *Stopword*. Setiap kata yang ada didalam *tweet* dipecah menjadi bagian-bagian tertentu dan apabila ada kata yang terdapat dalam daftar *Stopword* maka kata tersebut akan dihapus. *Library* untuk tokenisasi dan *Stopword* yang digunakan pada penelitian ini adalah *library* yang sudah disediakan nltk (*natural language toolkit*) yang dapat digunakan pada data berbahasa Indonesia atau bahasa inggris. Pembuangan *Stopword* ditujukan untuk menghapus kata tidak bermakna atau tidak relevan. Contoh hasil proses *Stopword removal* dapat dilihat pada TABEL VI.

TABEL VI
 CONTOH HASIL PENGHILANGAN STOPWORD BERBAHASA INDONESIA DAN INGGRIS

Train tweet	Stopword remove
@idextratime klo gw jd Ronaldo, emyu bakal gw beli setelah pildun	Idextratim klo gw ronaldo emyu gw beli pildun
Fix gausah masuk squad lagi. Mau squad banyak cedera mainin akademi aja,	fix gausah masuk squad squad cedera mainin akademi anggep desember

anggap desember emyu sedekah ke anak yatim januari tendang.	emyu sedekah anak yatim januari tending
He's credited Rashford for one aspect in particular #mufc	credit rashford one aspect particular mufc
Man of the match. Congratulations marcus	man match congratulation Marcus

D. *Training Klasifikasi*

Dataset yang sudah melalui tahap praproses, selanjutnya akan dilatih dengan menggunakan algoritme CNN dan *Naive Bayes*. Sebelum dilakukan proses *Training* pada algoritme CNN, tidak semua data *tweet* memiliki panjang data yang sama. Untuk menangani perbedaan panjang datanya pada penelitian ini ditentukan panjang maksimumnya adalah 1000. Model yang digunakan pada algoritme CNN adalah model *sequential*. Dalam konfigurasi, model *sequential* CNN memiliki beberapa layer konvolusi dan *pooling* yang bertujuan untuk mengambil fitur-fitur penting dari data input. Setelah itu, *vector* fitur diolah oleh layer *Dense* dengan fungsi aktivasi ReLU (*Rectified Linear Unit*), dan kemudian di-outputkan sebagai kelas menggunakan layer *Dense* dengan fungsi aktivasi sigmoid. Model ini yang digunakan pada proses *Training* dengan algoritme CNN untuk melakukan klasifikasi teks ke dalam kategori yang tepat. *Summary* pada model *sequential* dapat dilihat pada Gbr 3.

```

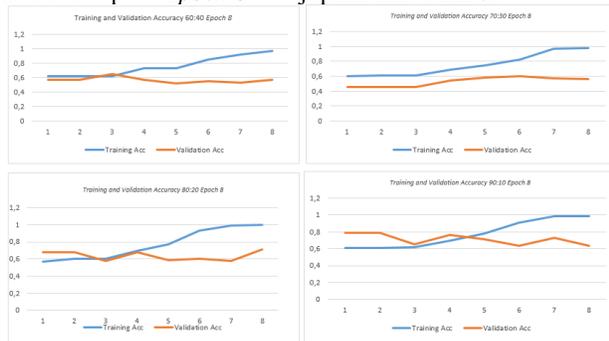
Model: "sequential"
-----
Layer (type)                Output Shape                Param #
-----
embedding (Embedding)       (None, 1000, 300)          1028700
conv1d (Conv1D)              (None, 1000, 128)          153728
max_pooling1d (MaxPooling1D) (None, 500, 128)           0
conv1d_1 (Conv1D)            (None, 500, 64)            32832
max_pooling1d_1 (MaxPooling1D) (None, 250, 64)           0
conv1d_2 (Conv1D)            (None, 250, 32)            8224
max_pooling1d_2 (MaxPooling1D) (None, 125, 32)           0
flatten (Flatten)            (None, 4000)                0
dense (Dense)                (None, 256)                 1024256
dense_1 (Dense)              (None, 1)                   257
-----
Total params: 2,247,997
Trainable params: 2,247,997
Non-trainable params: 0
    
```

Gbr 3. *Model Sequential Summary*

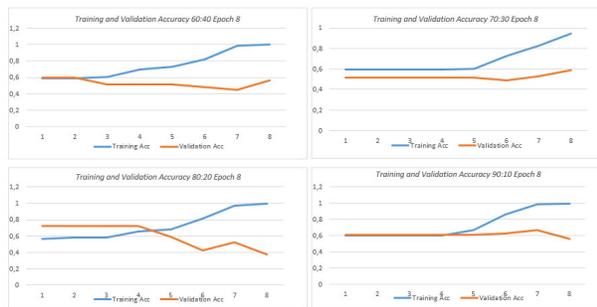
Pada penelitian ini model dari *Naive Bayes* yang digunakan adalah model *Bernoulli Naive Bayes*. Model *Bernoulli* ini menggunakan fitur *Boolean* (variabel biner) untuk nilai input. Model ini banyak digunakan untuk tugas klasifikasi dokumen dengan fitur kemunculan istilah biner digunakan untuk

melakukan klasifikasi kalimat yang ada di dalam suatu dokumen.

Pada algoritme CNN digunakan tiga *Epoch*, yaitu *Epoch 8*, *Epoch 10*, dan *Epoch 12*. Tidak ada jumlah *Epoch* yang optimal. Faktor utama yang mempengaruhi adalah kesalahan pada pelatihan dan validasi. Pada proses pelatihan ketika kedua set data pelatihan dan validasi dipelajari oleh model, dengan menggunakan grafik, dapat dilihat bahwa kesalahan pelatihan dan validasi terus menurun. Ini menunjukkan bahwa pembelajaran berhasil dan berjalan dengan baik. Namun, ketika kesalahan validasi meningkat, itulah saat proses pelatihan harus dihentikan. Ketika kesalahan validasi meningkat, itu berarti model mengalami *overfitting*, yang berarti model menghafal data tanpa benar-benar mempelajarinya [18]. Hasil uji coba klasifikasi pada *Epoch 8* tersaji pada Gbr 4 dan 5.

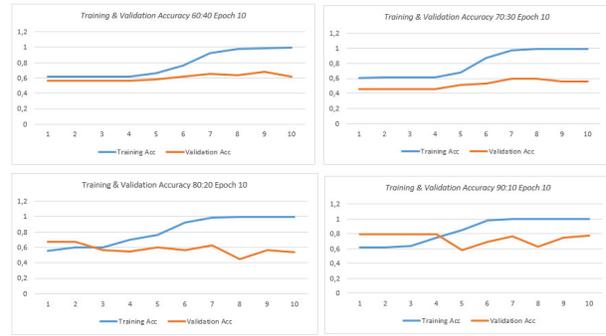


Gbr 4. Grafik Training CNN Dataset Berbahasa Inggris Epoch 8

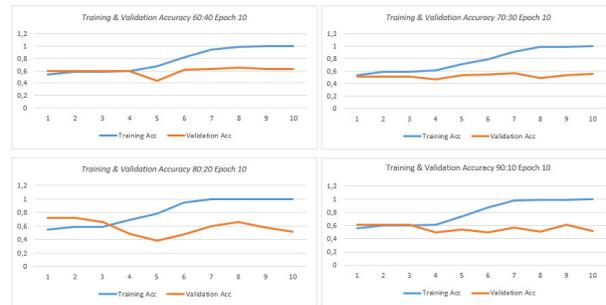


Gbr 5. Grafik Training CNN Dataset Berbahasa Indonesia Epoch 8

Dari 4 skenario *Training* yang telah dilakukan, pada *dataset* berbahasa Inggris didapatkan nilai akurasi *Training* terakhir dari masing-masing proporsi *data training* dan *data testing* adalah 0.9722, 0.9825, 0.9972, 0.9889. Sedangkan pada *dataset* berbahasa Indonesia nilai akurasi *Training* terakhir yang didapatkan adalah 0.9963, 0.9460, 0.9944, 0.9914. Dapat disimpulkan bahwa nilai akurasi *Training* dari skenario *Training* pada *Epoch 8* mengalami kenaikan dan penurunan. Kemudian nilai dari *Training Epoch 10* tersaji pada Gbr 6 dan 7.

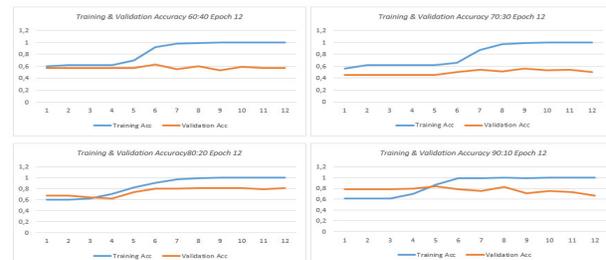


Gbr 6. Grafik Training CNN Dataset Berbahasa Inggris Epoch 10

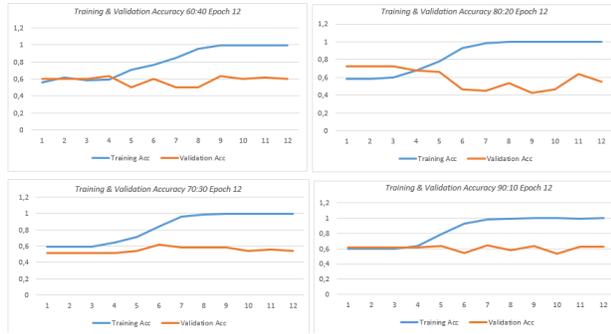


Gbr 7. Grafik Training CNN Dataset Berbahasa Indonesia Epoch 10

Pada 4 skenario *Training* yg dilakukan di *Epoch 10*, didapatkan nilai akurasi *Training* terakhir pada *dataset* berbahasa Inggris sebesar 0.9963, 0.9889, 0.9944, 0.9988, sedangkan pada *dataset* berbahasa Indonesia nilai akurasi *Training* terakhir yg diperoleh adalah 0.9963, 0.9968, 0.9972, 0.9988. Nilai akurasi *Training* pada *dataset* berbahasa Inggris mengalami kenaikan dan penurunan seperti yang terjadi pada *Epoch 8*, sedangkan pada *dataset* berbahasa Indonesia nilai akurasi *Training*nya terus naik dan tidak mengalami penurunan. Selanjutnya yang terakhir tahap *Training* di nilai *Epoch 12* dari *Convolutional Neural Network* tersaji pada Gbr 8 dan 9.



Gbr 8. Grafik Training CNN Dataset Berbahasa Inggris Epoch 12



Gbr 9. Grafik Training CNN Dataset Berbahasa Indonesia Epoch 12

Pada 4. skenario *Training* yg dilakukan di *Epoch* 12, didapatkan nilai akurasi *Training* terakhir pada *dataset* berbahasa Inggris adalah 0.9981, 0.9984, 0.9972, 0.9975, sedangkan pada *dataset* berbahasa Indonesia diperoleh nilai akurasi terakhir sebesar 0.9981, 0.9986, 0.9984, 0.9963. Pada nilai *Epoch* 12, nilai akurasi dari masing-masing *dataset* mengalami kenaikan dan penurunan. Dari Hasil *Training dataset* terdapat nilai *Training accuracy* dan *validation accuracy*. *Training accuracy* adalah nilai dari penghitungan akurasi dari *Training dataset* serta prediksi dari modelnya., sedangkan *validation accuracy* adalah hasil penghitungan ketepatan dari *validation dataset* serta prediksi dari modelnya dengan input informasi dari *validation dataset* tersebut. Nilai *Training accuracy* dapat digunakan untuk kemajuan model klasifikasi yang sudah dibuat dalam *Training*, dan nilai *validation accuracy* digunakan untuk mengukur kualitas model klasifikasi serta seberapa baik model klasifikasi digunakan untuk memprediksi data baru.

E. Evaluasi Hasil Klasifikasi

Pada tahap ini klasifikasi tweet yang sudah dibuat dievaluasi. Evaluasi klasifikasi tweet dilakukan dengan mengukur *accuracy*, *precision*, *recall*, dan *f1-score* dari hasil perhitungan algoritme CNN dan *Naïve Bayes*. Hasil perbandingan dari perhitungan klasifikasi pada dua metode tersebut diperlihatkan pada TABEL VII berikut.

TABEL VII
HASIL PERBANDINGAN DAN PERHITUNGAN KLASIFIKASI CNN DATASET BERBAHASA INGGRES

Epoch	Proposi Data (Data Latih : Data Uji)	CNN (Data Berbahasa Inggris)			
		A	P	R	F1
8	90 : 10	92%	91%	93%	92%
	80 : 20	90%	91%	88%	89%
	70 : 30	81%	80%	79%	79%
	60 : 40	77%	78%	73%	74%

10	90 : 10	94%	94%	93%	94%
	80 : 20	86%	85%	86%	85%
	70 : 30	81%	80%	78%	79%
	60 : 40	70%	69%	69%	69%
12	90 : 10	92%	91%	93%	92%
	80 : 20	93%	93%	91%	93%
	70 : 30	84%	85%	80%	82%
	60 : 40	70%	68%	68%	68%

TABEL VIII
HASIL PERBANDINGAN DAN PERHITUNGAN KLASIFIKASI CNN DATASET BERBAHASA INGGRES

Epoch	Proposi Data (Data Latih : Data Uji)	CNN (Data Berbahasa Indonesia)			
		A	P	R	F1
8	90 : 10	91%	91%	91%	91%
	80 : 20	81%	82%	82%	81%
	70 : 30	74%	73%	73%	73%
	60 : 40	65%	64%	64%	69%
10	90 : 10	90%	89%	90%	89%
	80 : 20	83%	83%	84%	83%
	70 : 30	75%	74%	73%	74%
	60 : 40	65%	64%	65%	64%
12	90 : 10	91%	91%	91%	91%
	80 : 20	84%	84%	83%	84%
	70 : 30	74%	74%	74%	74%
	60 : 40	65%	63%	55%	64%

TABEL IX
HASIL PERBANDINGAN DAN PERHITUNGAN KLASIFIKASI CNN DATASET BERBAHASA INGGRES DAN INDONESIA

Proposi Data (Data Latih : Data Uji)	Naïve Bayes (Data Berbahasa Inggris)				Naïve Bayes (Data Berbahasa Indonesia)			
	A	P	R	F1	A	P	R	F1
90 : 10	79%	87%	71%	72%	75%	85%	68%	68%
80 : 20	77%	86%	70%	71%	71%	83%	65%	64%
70 : 30	76%	86%	69%	69%	68%	82%	62%	59%
60 : 40	70%	83%	61%	59%	65%	81%	56%	49%

Hasil perhitungan klasifikasi metode CNN memiliki hasil nilai pengujian yang lebih baik secara keseluruhan daripada

metode *Naïve Bayes*. Nilai akurasi paling tinggi dari metode CNN dengan *dataset* berbahasa Inggris ada pada nilai *Epoch* 10 sebesar 94%, sedangkan pada metode CNN dengan *dataset* berbahasa Indonesia nilai akurasi paling tinggi ada pada nilai *Epoch* 12 yaitu sebesar 91%. Sedangkan pada metode *Naïve Bayes*, nilai akurasi paling tinggi dengan *dataset* berbahasa Inggris adalah 79%, dan pada *dataset* berbahasa Indonesia nilai akurasi tertingginya adalah sebesar 75%. Nilai *accuracy*, *presicion*, *recall*, dan *f1-score* yang telah didapatkan dari 2 metode tersebut memiliki rincian dan makna tersendiri. *Accuracy* adalah rasio antara prediksi yang benar (positif atau negatif) dan total data. *Precision* adalah rasio ketepatan antara label yang sebenarnya dengan label yang telah dilakukan oleh metode klasifikasi. *Recall* adalah rasio kemampuan pengklasifikasian untuk menemukan semua data yg benar. Terakhir, *f1-score* merupakan rata-rata antara presisi dan *recall* dengan skor terbaik yaitu 1.0 atau 100% dan skor terendah yaitu 0 atau 0%. Secara umum, *f1-score* memiliki nilai lebih rendah dibandingkan dengan nilai akurasi karena masuknya nilai presisi dan *recall* kedalam perhitungan *f1-score*.

Menurut penelitian *Convolutional neural networks for sentence classification* yang dilakukan oleh Yoon Kim, [19] hasil penelitiannya adalah model CNN lebih unggul dalam memperoleh akurasi lebih baik dari metode *Naïve Bayes*, terutama pada *dataset* yang lebih besar dan memiliki kompleksitas yang tinggi. Namun, model *Naïve Bayes* tetap menghasilkan performa yang baik pada *dataset* kecil dengan karakteristik yang sederhana. Model CNN mengatasi permasalahan seperti variasi kata dalam dokumen, sementara model *Naïve Bayes* lebih rentan terhadap permasalahan tersebut. Namun, model *Naïve Bayes* menghasilkan interpretasi hasil klasifikasi yang lebih mudah dipahami daripada model CNN [20]. Model CNN dapat mempelajari fitur-fitur yang berguna secara otomatis dari teks. Sementara model *Naïve Bayes* hanya bergantung pada matriks kemunculan kata dan frekuensi [21].

V. KESIMPULAN

Pada penelitian ini, Kesimpulan dari perbandingan kinerja CNN dan *Naive Bayes* pada analisis sentimen performa Manchester United di twitter ini adalah perbandingan hasil klasifikasi antara metode CNN dengan *Naive Bayes* memiliki nilai klasifikasi yang berbeda. Pada *dataset* berbahasa Inggris CNN lebih unggul dibandingkan *Naive Bayes* dengan nilai *accuracy* 94%, *precision* 94%, *recall* 93%, dan *f1-score* 94%, dan *Naive Bayes* dengan akurasi 79%, presisi 87%, *recall* 71%, dan *f1-score* 72%. Pada *dataset* berbahasa Indonesia CNN juga lebih unggul dibandingkan *Naive Bayes* dengan *accuracy* 91%, *precision* 91%, *recall* 91%, dan *f1-score* 91%, dan *Naive Bayes* dengan *accuracy* 75%, *precision* 85%, *recall* 68%, dan *f1-score* 68%.

Mengingat masih adanya kekurangan pada penelitian ini, penulis menyarankan pada penelitian selanjutnya Diharapkan untuk meningkatkan jumlah data tweet, yang menghasilkan lebih banyak *data training*, dan menyetarakan kelas data positif

dan negatif, sehingga sistem klasifikasi dapat mencapai nilai akurasi yang lebih baik. Khusus pada data berbahasa Indonesia dapat ditambahkan dengan perbaikan kata tidak baku, karena data tweet dari pengguna twitter di Indonesia masih banyak *tweet* yang bahasanya bercampur dengan bahasa Inggris atau bahasa daerah.

REFERENSI

- [1] Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. In *Computer Science Review*. <https://doi.org/10.1016/j.cosrev.2017.10.002>
- [2] Giulianotti, R. (2018). Supporters, followers, fans, and flâneurs: A taxonomy of spectator identities in football. In *Sport, Power, and Society: Institutions and Practices: A Reader*. <https://doi.org/10.4324/9780429497216>.
- [3] Sigalingging, B. P. H. (2018). *Red Devils the Legend: Sejarah Manchester United dari Masa ke Masa*. Media Pressindo.
- [4] Ratnawati, F., & Winarko, E. (2018). Sentiment Analysis of Movie Opinion in Twitter Using Dynamic Convolutional Neural Network Algorithm. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 12(1), 1.
- [5] Falahah, & Nur, D. D. A. (2015). Pengembangan Aplikasi Sentiment Analysis Menggunakan Metode *Naive Bayes*. *Seminar Nasional Sistem Informasi Indonesia, November*, 335–340.
- [6] Ramdhani, M. A., Ramdhani, M. A., Maylawati, D. S. Adillah, & Mantoro, T. (2020). Indonesian news classification using convolutional neural network. *Indonesian Journal of Electrical Engineering and Computer Science*, 19(2), 1000–1009.
- [7] Eka Putra, W. S. (2016). Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101. *Jurnal Teknik ITS*, 5(1). <https://doi.org/10.12962/j23373539.v5i1.15696>
- [8] Kim, Y. (2014). Convolutional neural networks for sentence classification. *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, 1746–1751. <https://doi.org/10.3115/v1/d14-1181>
- [9] Feldman, R., Sanger, J., & Press, C. U. (2007). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press. https://books.google.co.id/books?id=U3EA_zX3ZwEC
- [10] Falahah, & Nur, D. D. A. (2015). Pengembangan Aplikasi Sentiment Analysis Menggunakan Metode *Naive Bayes*. *Seminar Nasional Sistem Informasi Indonesia, November*, 335–340.
- [11] Singh, G., Kumar, B., Gaur, L., dan Tyagi, A. 2019. Comparison between Multinomial and Bernoulli *Naive Bayes* for Text Classification. In *2019 International Conference on Automation, Computational and Technology Management, ICACTM 2019*. hal. 593–596. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ICACTM.2019.8776800>
- [12] Feldman, R., Sanger, J., & Press, C. U. (2007). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press. https://books.google.co.id/books?id=U3EA_zX3ZwEC
- [13] Berry, M. W., & Kogan, J. (2010). Text Mining: Applications and Theory. In *Text Mining: Applications and Theory*. <https://doi.org/10.1002/9780470689646>
- [14] Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*. <https://doi.org/10.2200/S00416ED1V01Y201204HLT016>
- [15] Bilton, N. (2013). *Hatching Twitter: A True Story of Money, Power, Friendship and Betrayal*. Hodder & Stoughton. <https://books.google.co.id/books?id=jXgPAAAAQBAJ>
- [16] Blumrodt, J (2014). *Enhancing Football Brands' Brand Equity*. *The Journal of applied Business Research*. 1551-1558.
- [17] Eliane Birba, D. (2020). A Comparative study of data splitting algorithms for machine learning model selection. *Degree Project in Computer Science and Engineering*, 1–23.

- [18] Afaq, S., & Rao, S. (2020). Significance Of *Epochs* On Training A Neural Network. *International Journal of Scientific & Technology Research*, 9(06), 485–488.
- [19] Kim, Y. (2014). Convolutional neural networks for sentence classification. *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, 1746–1751. <https://doi.org/10.3115/v1/d14-1181>
- [20] Zhang, Y., & Wallace, B. (2015). A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification. <http://arxiv.org/abs/1510.03820>
- [21] Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural language processing (almost) from scratch. *Journal of Machine Learning Research*.