ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 15, Nomor 1, Tahun 2026, Halaman 46 - 56
Online di: https://ejournal3.undip.ac.id/index.php/gaussian/

KLASIFIKASI SENTIMEN KASUS *ONLINE TRADING* BINOMO PADA TWITTER MENGGUNAKAN ALGORITMA *CONVOLUTIONAL NEURAL NETWORK* (CNN)

Elva Nadia Nugroho[1*], Tatik Widiharih[2], Arief Rachman Hakim[3]
[1,2,3] Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro
***e-mail**: elvanadian@gmail.com

DOI: 10.14710/j.gauss.15.1.46-56

Article Info:
Received: 2024-12-07
Accepted: 2025-12-30
Available Online: 2026-04-29

Keywords:
Binomo case; Twitter; Sentiment Analysis; Convolutional Neural Network; Word2Vec

Abstract: Investment in futures trading and technology is growing in Indonesia, making many domain sites of online trading companies appear that are easy for everyone to access. The Binomo app is so viral among the public because of an ad that displays professional traders who can earn 1000 USD a day without leaving their homes. Binary options trading ultimately causes losses for some people, allegedly caused by affiliates. The Binomo case is widely discussed on social media, especially Twitter. Tweets displaying public opinion on Twitter can be used for sentiment analysis and categorizing public opinion on the Binomo case. 2.686 tweets were collected by Twitter scraping between January 1 and April 1, 2022. 1.519 tweets were left after pre-processing. The data were processed using the *Convolutional Neural Network* algorithm with the *Word2Vec* method to determine their accuracy and identify topics often discussed by the public on Twitter. A CNN model with 70% training data, 3, 4, 5, kernel sizes, 4 batch sizes, 30 epochs, and Adam optimizers was used to build the classification in this research. The accuracy value obtained from the performance evaluation of the *Convolutional Neural Networ* model research was 89%.

1. PENDAHULUAN

Perkembangan investasi di bidang perdagangan berjangka dan teknologi di Indonesia menyebabkan banyaknya situs *online trading* yang muncul dan mudah diakses oleh semua orang. *Binary option* adalah salah satu *online trading* yang semakin populer di Indonesia (Dirgantara, 2021). Aplikasi *binary option* Binomo begitu viral di kalangan masyarakat karena iklan yang menunjukkan seorang *trader* profesional yang dapat menghasilkan 1.000 USD setiap hari tanpa meninggalkan rumah. *Trading binary option* pada akhirnya menyebabkan kerugian bagi beberapa orang diduga disebabkan oleh para *affiliator*-nya. Nama Doni Salmanan menjadi topik hangat di media sosial setelah Indra Kenz, seorang *influencer* yang mempopulerkan aplikasi Binomo ditetapkan sebagai tersangka kasus investasi bodong menggunakan Binomo. Doni kemudian dipanggil oleh pihak berwajib soal kasus serupa dengan Indra, tetapi dengan *platform trading* Quotex (Orlando, 2022).

Kasus Binomo ramai diperbincangkan di media sosial terutama Twitter. Twitter menjadi media sosial paling populer di kalangan pengguna internet. *Tweet* berisikan cuitan pengguna akun Twitter mengenai suatu topik. *Tweet* ini dapat digunakan dalam analisis sentimen dan pengklasifikasian opini masyarakat terkait kasus *online trading* Binomo. Algoritma *deep learning* yang biasa digunakan untuk menganalisis data berupa teks adalah *Convolution al Neural Network* (CNN), *Recurrent Neural Network* (RNN), *Long Short Term Memory* (LSTM), dan *Gated Recurrent Unit* (GRU) (Zhang, *et al*., 2018).

Berdasarkan penelitian yang dilakukan oleh Kim (2014), model CNN mendapatkan hasil yang baik untuk permasalahan klasifikasi sentimen dari data tekstual. Tingkat akurasi yang diperoleh penelitian Kim dengan *dataset* yang dinamai MPQA menghasilkan akurasi tinggi sebesar 89,6%. Algoritma CNN dianggap memiliki kemampuan ekstraksi fitur yang

sangat kuat dan sesuai untuk analisis opini artikel yang cukup panjang (Rhanoui, *et al*., 2019). Berdasarkan penelitian yang dilakukan Ilmania dan Cahyawijaya (2018) disimpulkan bahwa penerapan *word embedding* untuk membuat *similar matrix* dapat meningkatkan akurasi. Percobaan ini menggunakan Word2Vec, yang memiliki ukuran kata 500 dimensi dan telah dilatih pada 27 juta kalimat dari data berita Bahasa Indonesia.

Penelitian Juwiantho, *et al*. (2020), "*Sentiment Analysis* Twitter Bahasa Indonesia Berbasis Word2Vec Menggunakan *Deep Convolutional Neural Network*" menggunakan model Word2Vec untuk representasi kata vektor Bahasa Indonesia. Penelitian yang dilakukan menunjukkan bahwa penggunaan Word2Vec meningkatkan proses pelatihan dan meningkatkan akurasi algoritma *Deep Convolutional Neural Network*, dengan akurasi sebesar 76,40%.

Berdasarkan latar belakang yang telah diuraikan, penelitian ini menggunakan klasifikasi sentimen dengan CNN dan Word2Vec yang diharapkan dapat membantu menilai opini publik terhadap kasus *online trading* Binomo di Indonesia secara lebih tepat. Pengambilan data pada penelitian ini dilakukan dengan *Twitter Scraping*. Penelitian ini diharapkan dapat menentukan *accuracy, precision*, dan *recall* yang dihasilkan dari klasifikasi sentimen kasus *online trading* Binomo di Indonesia menggunakan algoritma *Convolutional Neural Network* dan mengidentifikasi topik yang sering dibicarakan dalam sentimen positif, negatif, dan netral terkait kasus *online trading* Binomo di Indonesia.

2. TINJAUAN PUSTAKA

Online trading adalah suatu sistem yang melibatkan *platform* investasi yang dilakukan secara online. *Binary option* adalah produk keuangan dimana pihak-pihak terlibat dalam transaksi yang didasarkan pada pemilihan opsi apakah aset akan kadaluwarsa atau tidak, berdasarkan harga aset tersebut. *Binary option* menawarkan "ya" atau "tidak". Dengan kata lain, sistem memprediksi atau menebak apakah aset akan naik atau turun dalam jangka waktu tertentu (Downey, 2021).

Salah satu produk yang dihasilkan oleh *platform* opsi biner ini adalah aplikasi Binomo. Binomo memiliki beberapa permasalahan yang signifikan untuk dibahas. Binomo menyediakan media untuk berjudi dengan melakukan aksi tebak-tebakan apakah harga suatu aset naik atau turun. Jika tebakan benar maka *trader* akan menerima uang dari *affiliator* dan sebaliknya jika salah maka uang akan hilang (Ariesta, 2022).

Twitter adalah layanan jejaring sosial yang memungkinkan penggunanya untuk mengirim dan membaca pesan berbasis teks hingga 280 karakter yang dikenal dengan sebutan kicauan (*tweet*). Pengguna bisa menulis *tweet* pada halaman beranda, bertukar pesan singkat melalui *direct message,* dan ikut berkomentar pada topik yang sedang *trending* di negara tertentu. *Scraping* adalah metode untuk mendapatkan informasi dari situs web secara otomatis tanpa menyalin secara manual (Ayani, *et al*., 2019). Tujuan dari teknik ini adalah untuk menemukan informasi pada bagian tertentu dari situs web.

Analisis sentimen, juga disebut *opinion mining*, adalah analisis mengenai pendapat, sentimen, evaluasi, penilaian, sikap, dan emosi individu terhadap entitas seperti produk, layanan, organisasi, individu, isu, peristiwa, dan topik lainnya (Liu, 2012). Analisis sentimen juga dapat menganalisis data teks seperti *tweet*, ulasan di media sosial, dan web (Pang dan Lee, 2008). *Text mining* adalah metode penggalian data teks yang semakin populer.

Text mining adalah penemuan informasi baru dan tidak diketahui sebelumnya oleh komputer dengan otomatis mengekstrak informasi dari berbagai sumber. Menggabungkan informasi yang berhasil diekstraksi dari berbagai sumber merupakan bagian penting dari

proses ini (Hearst, 2003). Kombinasi *data mining* dan *text mining* dapat membantu mengatasi masalah klasifikasi, klastering, dan prediksi data tekstual.

Transformasi data tekstual menjadi format yang lebih sederhana agar mudah dibaca oleh sistem sebagai persiapan untuk pemrosesan selanjutnya disebut sebagai *text pre-processing.* Proses *pre-processing* teks untuk penelitian ini mencakup *case folding*, menghapus URL, menghapus *escape* HTML, menghapus *mention*, menghapus *punctuation*, menghapus angka, menghapus duplikasi, dan normalisasi kata. Proses memotong dokumen menjadi potongan-potongan kecil dan disertai dengan proses untuk menghilangkan karakter tertentu, seperti tanda baca disebut *tokenizing.* Untuk mengoptimalkan analisis, karakter tanda baca dihilangkan. Kalimat dibagi menjadi kumpulan kata menggunakan *delimiter* spasi.

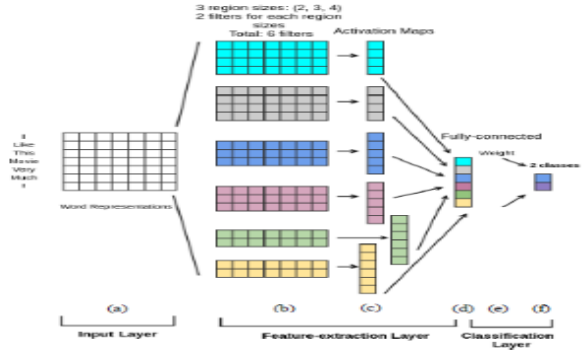
Sentiment scoring merupakan proses yang digunakan untuk melabelkan suatu pernyataan sehingga dapat digolongkan menjadi sentimen positif atau negatif berdasarkan kamus sentimen. Menurut Koto dan Rahmaningtyas (2018), kamus *Indonesian Sentiment Lexicon* (InSet) mengandung 6.609 kata negatif dan 3.609 kata positif, masing-masing dengan skor antara -5 dan +5.

Stopwords removal digunakan untuk menghilangkan kata-kata yang tidak penting atau tidak memberikan pesan dan kesan yang signifikan dalam teks. Penggunaan *stopwords removal* ditujukan untuk meningkatkan akurasi sistem klasifikasi sentimen (Ghag dan Shah, 2015). Proses *stemming* mengubah kata berimbuhan menjadi kata dasar dengan membuang sisipan, imbuhan, dan akhiran.

Word vector representation merupakan bagian dari algoritma *deep learning*. Setiap kata memiliki vektor yang mewakili makna dari kata tersebut dengan dimensi vektor yang bervariasi. *Word vector representation* juga disebut sebagai *word embedding*. Word2Vec adalah *word vector representation* yang dibuat oleh Google. Metode ini diciptakan oleh Tomas Mikolov. Word2Vec merepresentasikan teks dalam bentuk *dense vector.*

Convolutional Neural Network (CNN), salah satu metode *deep learning* yang berasal dari pengembangan *Multi-Layer Perceptron* (MLP), dirancang untuk mengolah data dua dimensi. CNN termasuk dalam kategori *Deep Neural Network.* CNN menggunakan dua metode *deep learning*, yaitu *feed forward* untuk klasifikasi dan *backpropagation* untuk tahap pembelajaran. Cara kerja CNN seperti yang ditunjukkan pada Gambar 2 menunjukkan bahwa CNN menggunakan dua metode, yaitu feedforward untuk klasifikasi dan backpropagation

Convolutional Neural Network (CNN) merupakan salah satu metode *deep learning* dari pengembangan *Multi Layer Perceptron* (MLP) yang didesain untuk mengolah data dua dimensi. CNN termasuk dalam jenis *Deep Neural Network* karena memiliki tingkat jaringan yang dalam. CNN memiliki dua metode yaitu klasifikasi menggunakan *feed forward* dan tahap pembelajaran menggunakan *backpropagation*. Cara kerja CNN mirip dengan MLP, seperti yang ditunjukkan pada Gambar 1, tetapi neuron di CNN digambarkan dalam dua dimensi, sedangkan pada MLP, setiap neuron hanya berukuran satu dimensi (Nasichuddin, *et al*., 2018).



Gambar 1. Arsitektur CNN (Nasichuddin, *et al*., 2018)

Pada awalnya, CNN dibuat untuk mengidentifikasi gambar, tetapi sekarang menjadi model yang dapat digunakan untuk berbagai tugas. Arsitektur CNN terdiri dari tiga lapisan, yaitu *input layer, feature extraction layer*, dan *clasification layer*. Arsitektur CNN memiliki beberapa parameter, seperti ukuran wilayah filter, jumlah filter, fungsi aktivasi, dan bentuk regularisasi. Lapisan utama pada arsitektur CNN terdiri dari *convolutional layer*, *pooling layer,* dan *fully connected layer.* (Nasichuddin, *et al*., 2018).

Berikut penjelasan untuk setiap *layer* pada CNN:

1. *Input layer*

Setiap kata dalam dokumen dipetakan dalam matriks. *Cell* dengan nilai 1 menunjukkan posisi kata, sedangkan *cell* tanpa kata diwakili dengan nilai 0.

2. *Feature-extraction Layer*

Lapisan lapisan yang terdapat dalam *feature* mentranslasikan suatu *input* menjadi *feature* berdasarkan ciri dari *input* yang terbentuk dari angka dalam vektor. Lapisan ini terdiri dari *convolutional layer* dan *pooling layer*.

a. Convolutional layer

Convolutional layer menghitung *output* neuron yang terhubung ke daerah lokal *input*. Setiap neuron menghitung produk titik antara bobotnya dan wilayah kecil yang terhubung ke volume *input*. Semua vektor kata ulasan yang ada dilakukan proses konvolusi pada lapisan ini. Ulasan dapat digambarkan sebagai berikut jika vektor-vektor kata digabungkan (Juwiantho, *et al*., 2020):

$$x_{i:n} = x_1 \oplus x_2 \oplus x_3 \oplus \ldots \oplus x_n \qquad (1)$$

Keterangan:

n : banyaknya vektor kata

x_i : vektor kata yang berada pada R dengan indeks ke $-\ i$

dengan $\oplus$ merupakan operator penggabungan. Secara umum, $x_{i:i+j}$ berarti hasil gabungan vektor-vektor kata dari indeks ke $-\ i$ sampai ke $-\ i+j$ seperti berikut:

$$x_{i:i+j} = x_i,\ x_{i+1},\ x_{i+2},\ \ldots,\ x_{i+j} \qquad (2)$$

Activation function kemudian memproses hasil *convolutional layer*. Fungsi aktivasi yang digunakan adalah fungsi aktivasi *ReLU* dan *Tanh*. *Rectified Linear Unit* (ReLU) digunakan sebagai fungsi non linier dalam penelitian ini. Fungsi ini memberikan batasan keluaran bernilai positif. Fungsi dapat dituliskan sebagai berikut:

$$ReLu_x = \begin{cases} 0\ if, & x < 0 \\ x\ if, & x \geq 0 \end{cases} \qquad (3)$$

Sehingga persamaan menjadi seperti berikut (Kim, 2014):

$$C_i = ReLU(net) \qquad (4)$$

$$net\ =\ w \cdot x_{i:i+h-1}$$

Untuk setiap *window* kata di dalam kalimat ulasan $\{x_{1:h}, x_{2:h}, x_{3:h}, \ldots, x_{n-h+1:n}\}$ diterapkan filter *w*, yang menghasilkan *feature map* berikut:

$$C = [C_1, C_2, C_3, \ldots, C_{n-h+1}] \quad (5)$$

b. *Pooling layer*

Pooling layer adalah lapisan yang mengurangi dimensi *feature map*, juga disebut sebagai langkah untuk *downsampling*, memungkinkan untuk mengatasi *overfitting* dan mempercepat komputasi karena parameter yang harus diperbaharui semakin sedikit.

Operasi *pooling* diterapkan pada setiap *feature map* yang diperoleh dari *convolution layer*. Nilai maksimum dari tiap *feature map* diambil pada lapisan ini (Kim, 2014).

$$\hat{C} = max\{C\} \quad (6)$$

$\hat{C}$ merupakan nilai maksimum dari *feature map* C. Vektor yang dihasilkan dari *pooling layer* adalah sebuah vektor yang terdiri dari nilai maksimum setiap *feature map* serta berjumlah *m* elemen.

$$Z = [\hat{C}_1, \hat{C}_2, \hat{C}_3, \ldots, \hat{C}_m] \quad (7)$$

Z merupakan vektor hasil dari *pooling layer* yang akan masuk kelapisan berikutnya.

3. *Classification Layer*

Lapisan ini berfungsi untuk mengklasifikasikan tiap neuron yang telah diekstrak pada fitur sebelumnya. *Classification layer* terdiri dari:

a. Flattern

Membentuk ulang fitur (*reshape feature map*) menjadi sebuah vektor agar bisa digunakan sebagai *input* dari *fully connected layer.*

b. Fully Connected (FC)

Fully connected layer dikenal juga sebagai *multilayer perceptron* (*feed-forward neural network*) yang berada pada ujung akhir CNN setelah mendapatkan hasil dari lapisan-lapisan sebelumnya, yaitu *convolution*, *activation* dan *pooling layer* yang menjadi sebagai masukannya berupa vektor satu dimensi. Lapisan ini akan mengolah data sehingga dapat menentukan hasil klasifikasi berdasarkan data yang telah dimasukkan.

c. *Output layer*

Output layer memiliki dua jenis fungsi yaitu *sigmoid* dan *softmax*. Fungsi sigmoid biasanya digunakan untuk mengaktifkan fungsi *output* dalam klasifikasi biner setelah mengambil nilai bilangan *real* dan mengembalikan nilai dalam rentang 0 hingga 1. Persamaan *sigmoid* dijelaskan sebagai berikut:

$$f(x_i) = \frac{1}{1+e^{(-x_j)}} \quad (8)$$

Softmax function menghitung probabilitas dari setiap kelas target atas semua kelas target, yang memungkinkan untuk menentukan kelas target untuk input. Karena fungsi ini mengembalikan probabilitas dari masing-masing kelas dan kelas target akan memiliki probabilitas yang lebih tinggi, maka fungsi ini cocok untuk digunakan dalam *multiclass clasification*. Rentang probabilitas output berkisar antara 0 dan 1, dan jumlah semua probabilitas akan sama dengan 1. Persamaan *softmax* dijelaskan sebagai berikut

$$f(x_i) = \frac{e^{(x_i)}}{\sum_{j=0}^{k} e^{(x_j)}} \quad (9)$$

Fungsi softmax menghasilkan vektor yang mewakili distribusi probabilitas dari daftar hasil potensial. Berdasarkan hasil probabilitas fungsi *softmax*, penentuan kelas dilihat berdasarkan nilai probabilitas tertinggi. Apabila probabilitas tertinggi berada di indeks ke – 0 maka data bernilai positif, apabila probabilitas tertinggi berada di indeks ke – 1 maka data bernilai negatif, dan apabila probabilitas tertinggi berada di indeks ke – 2 maka data bernilai netral.

Evaluasi kinerja klasifikasi diperlukan guna mengetahui bagaimana performa model yang dikembangkan. Pada penelitian ini, metode evaluasi *confusion matrix* digunakan untuk mengukur akurasi. Pengukuran dilakukan dengan menilai tingkat akurasi dari prediksi klasifikasi yang dihasilkan.

Tabel 1. *Confusion Matrix Multiclass*

		Prediksi		
		Positif	Negatif	Netral
Aktual	Positif	TPos	FPosNeg	FPosNet
	Negatif	FNegPos	TNeg	FNegNet
	Netral	FNetPos	FNetNeg	TNet

McNaught dan Lam (2010) berargumen bahwa visualisasi *wordcloud* dari teks akan memudahkan pengamat dalam melihat gagasan dan pendirian sang penulis teks sehingga dapat menjadi alat bantu dalam melakukan analisis terhadap sebuah wacana tertulis.

3. METODE PENELITIAN

Data yang digunakan pada penelitian ini merupakan data kualitatif yang diperoleh dari proses *twitter scraping* menggunakan *library snscrape* pada tanggal 1 Januari 2022 – 1 April 2022 dengan kata kunci "kasus *online trading* binomo". *Tweet* yang didapatkan sebanyak 2.686 *tweets* dan berbahasa Indonesia. Variabel terikat pada penelitian ini adalah sentimen masyarakat mengenai kasus *online trading* binomo dan variabel bebas adalah opini masyarakat terhadap kasus *online trading* binomo pada media sosial Twitter yang ditunjukkan pada *tweet*. Data *tweet* yang diperoleh dibersihkan dari duplikat hingga tersisa 1.159 *tweets*.

Langkah-langkah analisis yang dilakukan pada penelitian ini adalah sebagai berikut:

1. *Twitter Scraping*
2. *Data Pre-processing*
3. *Tokenizing*
4. *Sentiment Scoring*
5. *Stopwords removal*
6. *Stemming*
7. *Word2Vec*
8. Klasifikasi dengan algoritma *Convolutional Neural Network*.
   a. Membagi data menjadi data latih (*training*) dan data uji (*testing*)
   b. Membangun model *Convolutional Neural Network*
   c. Mengevaluasi hasil kinerja klasifikasi menggunakan *confusion matrix*
9. Interpretasi dalam bentuk visual menggunakan *wordcloud*

4. HASIL DAN PEMBAHASAN

Pengumpulan data *tweets* pengguna aplikasi Twitter dilakukan menggunakan *library snscrape* pada *Google Colaboratory* yang diambil dari tanggal 1 Januari 2022 – 1 April 2022 dan diperoleh sebanyak 2.686 *tweets*. Proses pengumpulan data dilanjutkan dengan menghilangkan duplikat dan *tweets* yang tidak memiliki arti hingga berkurang menjadi 1.519 *tweets*. Data *tweets* yang akan diolah disimpan dalam format CSV (*Comma Separated Value*).

Tahapan *pre-processing* yang dilakukan pada penelitian ini adalah sebagai berikut:

1. *Case Folding*
2. *Remove* URL
3. *Unescape* HTML

4. *Remove Mention*
5. *Remove Punctuation*
6. *Remove Number*
7. *Remove Duplicate*
8. Normalisasi kata

Pada proses *tokenizing*, *tweet* akan dipisah menjadi potongan kata dengan acuan pemisah berupa spasi. Proses ini dilakukan untuk memudahkan pelabelan data dalam *sentiment scoring*. Pada penelitian ini, data dilabelkan menggunakan kamus *Indonesian Sentiment Lexicon* (InSet). Hasil pelabelan data yang dilakukan program diperoleh 89 *tweets* berlabel positif, 1.383 *tweet* berlabel negatif, dan 47 *tweets* berlabel netral.

Stopwords removal bertujuan untuk menghilangkan kata-kata yang bersifat tidak penting atau tidak memberikan kesan atau pesan secara signifikan dalam *tweets*. *Stemming* dilakukan untuk mengubah kata berimbuhan menjadi kata dasar dengan menghapus sisipan, imbuhan dan akhiran.

Word embedding bertujuan untuk mengubah suatu kata menjadi vektor. Vektor yang dihasilkan dari proses ini akan merepresentasikan dari setiap kata yang ada dengan memperhatikan susunan kata pada *tweet*. Pada penelitian ini proses *word embedding* menggunakan *library gensim* dengan metode *Word2vec*. Data yang menjadi input yaitu kumpulan *token* dari hasil proses *text preprocessing*. Vektor yang dihasilkan pada proses ini memiliki panjang dimensi 100. Hasil sintaks dan *output* Word2Vec untuk kata binomo, korban, dan *trading* dapat dilihat pada Tabel 2.

Tabel 2. Hasil *Word2Vec*

Input Kata	Vektor
binomo	[0.06496565 -0.08004978 -0.071855515 ⋯ 0.04498074 -0.027405078 -0.103208244]
korban	[0.04252509 -0.05913359 -0.04796673 ⋯ 0.03542084 -0.018012797 -0.07736833]
trading	[0.03475038 -0.04999445 -0.04166746 ⋯ 0.02147232 -0.013331035 -0.055947356]

Data yang telah melalui tahapan *pre-processing* hingga *Word2Vec* dibagi menjadi data latih dan data uji dengan perbandingan 70%:30%.

Pembuatan arsitektur model CNN berdasarkan perancangan arsitektur yang telah dibuat sebelumnya yaitu di mulai dari data hasil *Word2Vec* yang masuk ke *input layer* akan dikalikan dengan kernel ukuran 3, 4, dan 5 untuk menghasilkan *feature map* merupakan proses dari *convolutional layer* dan masing-masing hasil tersebut dilakukan fungsi aktivasi *ReLu* sehingga tidak ada hasil berupa nilai negatif. Kemudian dilakukan operasi *pooling* dimana pada penelitian ini akan menggunakan *MaxPooling*. Hasil dari operasi *pooling* tersebut akan masuk ke bagian *fully connected.* Proses terakhir dilakukan proses *softmax* untuk menghasilkan probabilitas dari hasil klasifikasi.

Pembuatan model ini dengan cara melatih sistem dengan memberikan contoh *tweet* beserta labelnya sebagai bahan pembelajaran bagi model. *Training* ini menggunakan *epoch* atau perulangan sebanyak 30 dengan batch size 4. Selama proses *training* akan dimonitor nilai *validation loss* menggunakan *EarlyStopping()* dengan minimum perulangan sebanyak 5. Apabila nilai *validation loss* tidak terjadi peningkatan maka perulangan akan berhenti. Hasil sintaks dan *output* Model CNN dapat dilihat pada Tabel 3.

Tabel 3. Model CNN

```
Epoch 1/30
186/186 [==============================] - ETA: 0s - loss: 0.2467 - accuracy:
0.9180
Epoch 1: val_loss improved from inf to 0.21464, saving model to best_model.h5
```

Hasil *output* menunjukkan bahwa pada proses *training* diperoleh model terbaik pada epoch kesatu, nilai *loss validation* telah meningkat dari nilai tak terhingga (inf) menjadi 0.21464. Model terbaik yang telah disimpan selama pelatihan disimpan ke dalam file best_model.h5. Pemilihan model terbaik ditentukan berdasarkan model yang memiliki nilai *loss validation* terendah.

Setelah melatih model *Convolutional Neural Network* (CNN) untuk klasifikasi sentimen kasus *online trading* Binomo pada twitter, hasil klasifikasi dapat dilihat dengan memprediksi label pada data uji. Hasil klasifikasi model CNN pada data uji dapat dilihat pada Tabel 4.

Tabel 4. Hasil Klasifikasi Model CNN

Data ke-	Tweet	Label Prediksi	Probabilitas
455	buang uang kaya judi binomo	Negatif	0,922969
863	semangat korban binomo affiliator	Positif	0,988709
292	emosi banget kakak korban kalah binomo	Negatif	1

Berdasarkan data ke- 455 dan ke- 292, model memprediksi bahwa data yang diuji termasuk ke dalam sentimen negatif dengan tingkat kepercayaan sebesar 0.922969 dan 1. Sedangkan pada data ke- 863, model memprediksi bahwa data termasuk ke dalam sentimen positif dengan tingkat kepercayaan sebesar 0,988709. Tingkat kepercayaan ini mencerminkan seberapa yakin model dalam membuat prediksi tersebut. Semakin tinggi nilai probabilitas, semakin besar kepercayaan model dalam memprediksi data yang diuji.

Dalam penelitian ini, kinerja model klasifikasi yang telah diuji dan dibangun dengan data latih akan dievaluasi. Hasil dari *confusion matrix* untuk algoritma *Convolutional Neural Network* dapat dilihat pada tabel 5.

Tabel 5. *Confusion Matrix* algoritma *Convolutional Neural Network*

Kelas Aktual	Kelas Prediksi		
	Positif	Negatif	Netral
Positif	4	32	0
Negatif	2	402	2
Netral	0	14	0

$$Accuracy = \frac{4 + 402}{6 + 448 + 2} x100\% = 89\%$$

Hasil perhitungan *accuracy* model *Convolutional Neural Network* diperoleh akurasi sebesar 89%, nilai akurasi menunjukkan seberapa dekat prediksi model dengan nilai sebenarnya dari seluruh data uji, dengan nilai akurasi tersebut artinya 89% data dapat terklasifikasikan dengan benar oleh model.

Tabel 6. Perhitungan *Precision*

	Kelas Prediksi		
	Positif	Negatif	Netral
TP	4	402	0
FP	2	32 + 14	2
Precision	4 / (4 + 2)	402 / (402 + 46)	0

Berdasarkan Tabel 6, dapat diukur hasil *precision* secara manual sebagai berikut:

$$Precision = \frac{0{,}67 + 0{,}89 + 0}{3} x100\% = 52\%$$

Hasil perhitungan *precision* model *Convolutional Neural Network* diperoleh nilai sebesar 52%, nilai ini dipengaruhi oleh *false positive* (FP) yang artinya menunjukkan

seberapa banyak model memprediksi kelas positif dari data yang memang masuk ke kelas positif.

Tabel 7. Perhitungan *Recall*

	Kelas Prediksi		
	Positif	Negatif	Netral
TP	4	402	0
FN	32	2 + 2	14
Recall	4 / (4 + 32)	402 / (402 + 4)	0

Berdasarkan Tabel 30, dapat diukur hasil *recall* secara manual sebagai berikut:

$$Recall = \frac{0{,}11 + 0{,}99 + 0}{3} x100\% = 37\%$$

Hasil perhitungan *recall* model *Convolutional Neural Network* diperoleh nilai sebesar 37%, nilai ini dipengaruhi oleh *false negative* (FN) dan *true positive* (TP) yang berarti keberhasilan sistem untuk dapat mengenali kembali sampel positif pada data uji yaitu 37%.

Penelitian ini menggunakan *wordcloud* sebagai visualisasi data *tweets* dengan tujuan untuk mengetahui topik yang sering muncul pada setiap kelas sentimen. Gambar 10 menampilkan *wordcloud* berupa kumpulan kata atau topik yang sering muncul pada kelas sentimen positif.



Gambar 2. *Wordcloud Tweets* Sentimen Positif

Gambar 2 menunjukkan bahwa topik atau kata dengan sentimen positif yang sering muncul dalam *tweets* yaitu kata "judi", "investasi" dan "affiliator". Berdasarkan *tweet* berita yang diunggah oleh akun @editornews_id dengan *headline* "*Crazy Rich* Medan Indra Kenz Bakal Diperiksa Polisi Dalam Kasus Judi Online dan Promosikan Investasi Binomo https://t.co/zio4YLsbef", masyarakat banyak memberikan opini positif terkait dengan penangkapan **affiliator** (Indra Kenz) aplikasi **judi** Binomo yang berkedok **investasi**. Kata-kata lain pada *wordcloud* sentimen positif yang berukuran lebih kecil seperti "polisi", "korban", "lapor", dan "uang" menunjukkan bahwa kinerja polisi dalam menyelidiki kasus Binomo diapresiasi masyarakat dan diharapkan uang korban yang melapor dapat kembali.



Gambar 3. *Wordcloud Tweets* Sentimen Negatif

Gambar 3 menunjukkan bahwa topik atau kata dengan sentimen negatif yang sering muncul dalam tweets yaitu kata "judi", "korban", dan "uang" kemudian dilanjutkan dengan kata "rugi" dan "tipu". Berdasarkan *tweet* berita yang diunggah oleh akun @terkinidotid dengan *headline* "Binomo Resmi Diblokir, Bappebti: Hati-hati! Bukan trading Tapi Judi https://t.co/OxNv80Eawo #News #binomo #bappebtiblokir", masyarakat banyak memberikan opini negatif terkait dengan berita kasus Binomo yang ternyata adalah aplikasi **judi** online sehingga para **korban** merasa ditipu dan dirugikan karena **uang** mereka dirampas oleh affiliator Binomo.



Gambar 4. *Wordcloud Tweets* Sentimen Netral

Gambar 4 menunjukkan bahwa topik atau kata dengan sentimen netral yang sering muncul dalam tweets yaitu kata "judi", "main", dan "trading". Berdasarkan *tweet* berita yang diunggah oleh akun @djawanewsmedia dengan *headline* "8 Korban Laporkan Aplikasi Binomo Platform Judi Berkedok Trading Bikin Rugi Rp2,4 Miliar https://t.co/3jmtrmjfJ7", masyarakat memberikan opini netral terkait dengan **bermain judi** Binomo membuat rugi namun pemain **trading** baru sadar ketika uang mereka sudah habis. Mereka bermain judi atas kesadarannya sendiri sehingga seharusnya mereka sudah tahu konsekuensi yang harus ditanggung perbuatannya sendiri.

5. KESIMPULAN

Analisis dan pembahasan mengenai klasifikasi sentimen kasus *online trading* Binomo pada twitter menggunakan algoritma *Convolutional Neural Network* menghasilkan akurasi sebesar 89%. atau dari 1.519 data tweet sebanyak 1.352 data terklasifikasikan dengan benar. Hasil prediksi klasifikasi juga menghasilkan nilai *precision* sebesar 52%, dan nilai *recall* sebesar 37%. Berdasarkan nilai *accuracy*, *precision*, dan *recall* ini menunjukkan bahwa performa kinerja algoritma *Convolutional Neural Network* ini baik dan dapat diterapkan untuk mengklasifikasikan data tweet kasus online trading Binomo di Indonesia. Masyarakat Indonesia cenderung beropini negatif terhadap kasus *online trading* Binomo di Indonesia. Hasil interpretasi *Wordcloud* menunjukkan bahwa topik yang sering dibahas dalam sentimen negatif yaitu banyaknya korban yang tertipu oleh aplikasi Binomo yang ternyata adalah aplikasi judi *online*. Para korban merasa dirugikan karena uang mereka dirampas oleh affiliator Binomo.

DAFTAR PUSTAKA

Ariesta, A. 2022. *Fakta Kasus Binomo yang Menghebohkan, Korban Rugi hingga Rp2,4 Miliar. https://www.idxchannel.com/economics/ini-fakta-fakta-kasus-binomo-yang-menghebohkan-korban-rugi-hingga-rp24-miliar.*

Ayani, D. D., Pratiwi, H. S., dan Muhardi, H. 2019. *Implementasi Web Scraping untuk Pengambilan Data pada Situs Marketplace*. Jurnal Sistem Dan Teknologi Informasi (JUSTIN), 7(4), 257. https://doi.org/10.26418/justin.v7i4.30930.

Dirgantara, H. 2021. *BAPPEBTI Pastikan Transaksi Binary Option di Indonesia Ilegal*. https://investasi.kontan.co.id/news/bappebti-pastikan-transaksi-binary-option-ilegal-di-indonesia-ilegal.

Downey, L. 2021. *Binary Option*. https://www.investopedia.com/terms/b/binary-option.asp.

Ghag, K. V., dan Shah, K. 2015. *Comparative Analysis of Effect of Stopwords Removal on Sentiment Classification. International Conference on Computer, Communication and Control (IC4)*. India: Institute of Electrical and Electronics Engineers (IEEE).

Hearst, M. 2003. *What Is Text Mining?*. SIMS,UC Berkeley. https://people.ischool.berkeley.edu/~hearst/text-mining.html.

Ilmania, A., dan Cahyawijaya, S., 2018. *Aspect Detection and Sentiment Classification using Deep Neural Network for Indonesian Aspect-Based Sentiment Analysis*. 2018 International Conference on Asian Language Processing (IALP), hal.62– 67.

Juwiantho, H., Setiawan, E. I., Santoso, J., dan Purnomo, M. H. (2020). *Sentiment Analysis Twitter Bahasa Indonesia Berbasis WORD2VEC Menggunakan Deep Convolutional Neural Network*. Jurnal Teknologi Informasi Dan Ilmu Komputer, *7*(1), 181–188. Retrieved from https://jtiik.ub.ac.id/index.php/jtiik/article/view/1758.

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.

Koto, F., dan Rahmaningtyas, G. Y. 2017. *Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs*. 2017 International Conference on Asian Language Processing (IALP), pp. 391-394. https://doi.org /10.1109/IALP.2017.8300625.

Liu, B. 2012. *Sentiment Analysis and Opinion Mining*. San Rafael: Morgan dan Claypool Publishers, LLC.

McNaught, C., dan Lam, P. 2010. *Using Wordle as a Supplementary Research Tool*. The Qualitative Report, 15(3), 630-643. https://doi.org/10.46743/2160-3715/2010.1167.

Nasichuddin, M. A., Adji, T. B., dan Widyawan, W., 2018. Performance Improvement Using CNN for Sentimen Analysis, *International Journal of Information Technology and Electrical Engineering 21,* 9–114.

Orlando, I. G. 2022. *Mengenal Apa Itu Quotex, Trading Judi Online di Kasus Doni Salmanan*. https://finance.detik.com/fintech/d-5981797/mengenal-apa-itu-quotex-trading-judi-online-di-kasus-doni-salmanan.

Pang, B., dan Lee, L. 2008. *Opinion Mining and Sentiment Analysis*. Foundations and Trends in Information Retrieval, 2, 1-135. https://doi.org/10.1561/1500000011.

Rhanoui, M., Mikram, M., Yousfi, S., dan Barzali, S. 2019. *A CNN-BiLSTM Model for Document-Level Sentiment Analysis*. Machine Learning and Knowledge Extraction, 1(3), 832–847. https://doi.org/10.3390/make1030048.

Zhang, L., Wang, S., dan Liu, B. 2018. *Deep Learning for Sentiment Analysis : A Survey*. WIREs Data Mining and Knowledge Discovery, 8(4), e1253. https://doi.org/10.1002/widm.1253.