

**ANALISIS SENTIMEN KEBIJAKAN PEMBANGUNAN IBU
KOTA NUSANTARA DI MEDIA SOSIAL X MENGGUNAKAN
*BIDIRECTIONAL LONG SHORT-TERM MEMORY***

WANDA NUGRAHA



**PROGAM STUDI MATEMATIKA
SEKOLAH SAINS DATA MATEMATIKA DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**

@Hak cipta milik IPB University

IPB University



- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Analisis Sentimen Kebijakan Pembangunan Ibu Kota Nusantara di Media Sosial X Menggunakan *Bidirectional Long Short-Term Memory*” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2025

Wanda Nugraha
G5401211045

ABSTRAK

WANDA NUGRAHA. Analisis Sentimen Kebijakan Pembangunan Ibu Kota Nusantara di Media Sosial X Menggunakan *Bidirectional Longs Short Term Memory*. Dibimbing oleh MOCHAMAD TITO JULIANTO dan MOHAMAD KHOIRUN NAJIB.

Pembangunan Ibu Kota Nusantara (IKN) merupakan salah satu terobosan kebijakan pemerintah yang memicu beragam respons publik sehingga menjadi penting untuk memahami bagaimana pandangan publik terhadap kebijakan ini. Tujuan penelitian ini mencakup mengeksplorasi kecenderungan sentimen, membangun model analisis sentimen menggunakan metode *Bidirectional Long Short-Term Memory* (BiLSTM) yang dioptimalkan melalui *hyperparameter tuning*, dan melakukan evaluasi terhadap kinerja model sebelum dan sesudah *tuning*. Temuan dari proses eksplorasi distribusi sentimen menunjukkan dominasi sentimen negatif sebesar 52%, diikuti netral sebesar 34%, dan positif 14%. Model awal yang telah dibangun menghasilkan akurasi uji sebesar 80% dan rata-rata akurasi sebesar 82% pada *10-fold cross validation*. Meskipun akurasi uji tetap 80%, model hasil *tuning* mengalami peningkatan *recall* pada kelas negatif dan positif, yang penting dalam konteks analisis sentimen kebijakan publik. Rata-rata akurasi setelah *tuning* meningkat menjadi 84% pada *10-fold cross validation* menunjukkan model memiliki kemampuan generalisasi yang lebih baik.

Kata kunci: Analisis sentimen, Ibu Kota Nusantara, media sosial X, BiLSTM

ABSTRACT

WANDA NUGRAHA. Analysis of the Nusantara Capital City Development Policy on Social Media X using Bidirectional Long Short Term Memory. Supervised by MOCHAMAD TITO JULIANTO and MOHAMAD KHOIRUN NAJIB.

The development of the new capital city, Ibu Kota Nusantara (IKN) is a groundbreaking government policy that has triggered diverse public responses, making it crucial to understand how the public views this policy. The objectives of this study include exploring sentiment trends, building a sentiment analysis model using the Bidirectional Long Short-Term Memory (BiLSTM) method optimized through hyperparameter tuning, and evaluating the model's performance before and after tuning. Findings from the sentiment distribution exploration process indicate a dominance of negative sentiment at 52%, followed by neutral at 34%, and positive at 14%. The initial model that has been built produced a test accuracy of 80% and an average accuracy of 82% in 10-fold cross-validation. Although the test accuracy remained at 80%, the tuned model experienced increased recall in both the negative and positive classes, which is important in the context of public policy sentiment analysis. The average accuracy after tuning increased to 84% in 10-fold cross-validation, indicating the model has better generalization capabilities.

Keywords: BiLSTM, Nusantara Capital City, sentiment analysis, social media-X



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2025¹
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

ANALISIS SENTIMEN KEBIJAKAN PEMBANGUNAN IBU KOTA NUSANTARA DI MEDIA SOSIAL X MENGGUNAKAN *BIDIRECTIONAL LONG SHORT-TERM MEMORY*

WANDA NUGRAHA

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Matematika

**PROGAM STUDI MATEMATIKA
SEKOLAH SAINS DATA MATEMATIKA DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Penguji pada Ujian Skripsi:
Elis Khatizah S.Si., M.Si., Ph.D.

Judul Skripsi : Analisis Sentimen Kebijakan Pembangunan Ibu Kota Nusantara
di Media Sosial X Menggunakan *Bidirectional Long Short-Term
Memory*

Nama : Wanda Nugraha
NIM : G5401211045

Disetujui oleh

Pembimbing 1:
Mochamad Tito Julianto S.Si., M.Kom.



Pembimbing 2:
Mohamad Khoirun Najib S.Si., M.Mat.



Diketahui oleh

Ketua Program Studi:
Dr. Donny Citra Lesmana S.Si., M.Fin.Math.
NIP 197902272005011001



Tanggal Ujian: 21 Juli 2025

Tanggal Lulus:



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan September 2024 sampai bulan April 2025 ini ialah analisis sentimen, dengan judul “Analisis Sentimen Kebijakan Pembangunan Ibu Kota Nusantara di Media Sosial X Menggunakan *Bidirectional Long Short-Term Memory*”.

Terima kasih penulis ucapkan kepada:

1. Bapak Rahmat Gunawan, Ibu Rosita Dewi, dan Wildan selaku orang tua dan adik yang telah memberikan dukungan, kasih sayang, semangat, motivasi, dan doa kepada penulis,
2. Bapak Mochamad Tito Julianto S.Si., M.Kom selaku dosen pembimbing 1, Bapak Mohamad Khoirun Najib S.Si., M.Mat selaku dosen pembimbing 2, dan Ibu Elis Khatizah S.Si., M.Si., Ph.D selaku dosen penguji yang telah banyak memberikan ilmu, waktu, masukan, dan saran kepada penulis dalam penulisan dan perbaikan tugas akhir ini,
3. seluruh dosen dan asisten dosen di Program Studi S-1 Matematika SSMI IPB University yang telah memberikan ilmu kepada penulis,
4. seluruh staf Program Studi S-1 Matematika SSMI IPB University yang turut membantu kebutuhan administrasi penulis,
5. NPM 212121025 yang telah lama kebersamai dan membantu memberikan dukungan serta doa kepada penulis selama penyusunan karya ilmiah,
6. Kak Daryl dan Kak Muhamad Rifqi Musyaffa selaku mentor yang selalu membantu dan memberikan inspirasi selama masa perkuliahan,
7. sahabat terdekat penulis, yaitu Dave, Rifqi, Hadziq, Sulthan, Asrofil, Gilang, Syadalian, Abi, Raihan, Rhafy, Juna, Jourdy, Imam, Rendi yang telah berjuang bersama penulis selama masa perkuliahan,
8. seluruh grup Homomorfisma, Semhas January, WKWK Club, healing2, risaners uhuy!!, KKNT Karawang Digoyang, yang telah banyak memberikan kesan positif semasa perkuliahan,
9. seluruh pihak yang telah membantu penulis dalam menyelesaikan penulisan karya ilmiah yang tidak bisa disebutkan satu-satu.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2025

Wanda Nugraha

DAFTAR ISI

DAFTAR TABEL	viii
DAFTAR GAMBAR	viii
DAFTAR LAMPIRAN	ix
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan	2
II TINJAUAN PUSTAKA	3
2.1 Kebijakan Pembangunan Ibu Kota Nusantara (IKN)	3
2.2 Media Sosial X	3
2.3 <i>Natural Language Processing</i> (NLP)	4
2.4 Analisis Sentimen	4
2.5 Pelabelan Semi-Otomatis	4
2.6 VADER (<i>Valence Aware Dictionary and sEntiment Reasoner</i>)	5
2.7 Penanganan Data Tak Seimbang dengan Augmentasi Data Teks	5
2.8 <i>Feature Engineering</i> pada Data Teks	6
2.9 <i>Artificial Neural Network</i> (ANN)	8
2.10 Fungsi Aktivasi	11
2.11 <i>Recurrent Neural Network</i> (RNN)	13
2.12 <i>Bidirectional Long-Short Term Memory</i> (BiLSTM)	16
2.13 <i>Cross-Entropy Loss Function</i>	21
2.14 <i>Adam Optimizer</i>	22
2.15 <i>Hyperparameter Optimization</i> Berbasis <i>Bayesian Optimization</i>	22
2.16 <i>K-Fold Cross Validation</i>	24
2.17 Evaluasi Model Klasifikasi	25
III METODE	28
3.1 Sumber dan Jenis Data	28
3.2 Tahapan Penelitian	29
3.3 Lingkungan Pengembangan	34
IV HASIL DAN PEMBAHASAN	36
4.1 Pra-proses Data	36
4.2 Eksplorasi Data	40
4.3 Konstruksi Model Awal BiLSTM	43
4.4 Konstruksi Model Akhir	46
V SIMPULAN DAN SARAN	52
5.1 Simpulan	52
5.2 Saran	52
DAFTAR PUSTAKA	53
LAMPIRAN	58
RIWAYAT HIDUP	69



DAFTAR TABEL

1	Pemetaan <i>label encoding</i> dan konversi <i>one-hot encoding</i>	6
2	Ilustrasi hasil proses <i>padding</i> pada <i>feature engineering</i>	7
3	Kategori <i>confusion matrix</i> untuk klasifikasi tiga kelas	26
4	Ulasan data penelitian	28
5	Spesifikasi parameter setiap lapisan arsitektur BiLSTM	33
6	Contoh ulasan hasil proses <i>data cleaning</i>	36
7	Contoh ulasan hasil proses <i>case folding</i>	37
8	Contoh ulasan hasil proses <i>normalization</i>	37
9	Contoh ulasan hasil proses pelabelan semi-otomatis	38
10	Contoh ulasan hasil proses <i>tokenization</i>	39
11	Contoh hasil proses <i>stopwords removal</i>	39
12	Contoh hasil proses <i>stemming</i>	40
13	Contoh teks sentimen negatif	41
14	Contoh teks sentimen netral	41
15	Contoh teks sentimen positif	41
16	Contoh hasil augmentasi <i>back translation</i> dan <i>synonym replacement</i>	42
17	Hasil evaluasi performa model awal	45
18	Hasil 10-fold <i>cross validation</i> pada model awal	46
19	Ruang pencarian <i>hyperparameter</i> beserta nilai optimalnya	47
20	Hasil <i>hyperparameter tuning</i>	47
21	Hasil evaluasi performa model setelah <i>tuning</i>	49
22	Hasil 10-fold <i>cross validation</i> pada model hasil <i>tuning</i>	51

DAFTAR GAMBAR

1	Diagram ilustrasi proses <i>tokenization</i>	7
2	Diagram ilustrasi proses <i>word embedding</i>	8
3	Ilustrasi arsitektur ANN (Zhang et al. 2019)	8
4	Ilustrasi <i>perceptron</i> satu <i>node</i> (Zhang et al. 2019)	9
5	Ilustrasi <i>perceptron</i> dengan <i>m node</i> (Naumov 2017)	10
6	Ilustrasi <i>multilayer perceptron</i> dua <i>hidden layer</i> (Haykin 2009)	11
7	Fungsi aktivasi ReLU (Yun 2023)	12
8	Fungsi aktivasi sigmoid (Yun 2023)	12
9	Fungsi aktivasi tanh (Yun 2023)	13
10	Fungsi aktivasi <i>softmax</i> (Islam 2023)	13
11	Ilustrasi arsitektur RNN (Lubis et al. 2024)	14
12	Pemrosesan RNN (a) <i>many to many</i> (b) <i>many to one</i> (c) <i>one to many</i> (Lubis et al. 2024)	15
13	Ilustrasi arsitektur model LSTM (Zhang et al. 2018)	16
14	Ilustrasi proses kerja bagian <i>forget gate</i> (Olah 2015)	17
15	Ilustrasi proses kerja <i>input gate</i> dan <i>new memory cell</i> (Olah 2015)	18
16	Ilustrasi proses memperbarui <i>cell state</i> (Olah 2015)	19
17	Ilustrasi proses kerja bagian <i>output gate</i> (Olah 2015)	20
18	Ilustrasi proses <i>output</i> akhir model pada LSTM (Olah 2015)	20

19	Ilustrasi Arsitektur BiLSTM (Alghifari <i>et al.</i> 2022).	21
20	Penerapan 10-fold cross validation	25
21	Ulasan isi kamus <i>colloquial Indonesian lexicon</i>	29
22	Diagram alir pelabelan VADER	30
23	Diagram alir penelitian	35
24	Distribusi proporsi kelas sentimen	40
25	Distribusi proporsi kelas setelah augmentasi data	42
26	Grafik perubahan akurasi selama proses pelatihan	43
27	Grafik perubahan <i>loss</i> selama proses pelatihan	43
28	Digram <i>confusion matrix</i>	44
29	Perkembangan hasil <i>tuning</i> per percobaan	48
30	Diagram <i>confusion matrix</i> dan performa model setelah <i>tuning</i>	48
31	Diagram batang perubahan performa pada kelas negatif	49
32	Diagram batang perubahan performa pada kelas netral	50
33	Diagram batang perubahan performa pada kelas positif	50

DAFTAR LAMPIRAN

1	Data kumpulan <i>post</i> IKN	59
2	<i>Code data cleaning</i> (sebelum penghapusan tanda baca)	59
3	<i>Code</i> untuk proses <i>case folding</i>	59
4	<i>Code</i> untuk proses <i>normalization</i>	59
5	<i>Code</i> untuk proses translasi ke bahasa Inggris	59
6	<i>Code</i> untuk proses pelabelan VADER	60
7	<i>Code</i> untuk proses <i>cleaning</i> (penghapusan tanda baca) pada teks yang sudah melalui proses pelabelan semi-otomatis	60
8	<i>Code</i> untuk proses <i>tokenization</i>	60
9	<i>Code</i> untuk proses <i>removal stopwords</i>	61
10	<i>Code</i> untuk proses <i>stemming</i>	61
11	<i>Code</i> untuk proses menggabungkan token dalam <i>stemming</i> menjadi satu kalimat	61
12	<i>Code</i> untuk proses augmentasi data <i>synonym replacement</i>	61
13	<i>Code</i> untuk proses <i>feature engineering</i> (<i>label encoding</i> , <i>one hot encoding</i> , <i>tokenization</i> , dan <i>padding</i>)	62
14	<i>Code</i> untuk membangun model	62
15	<i>Compile model code</i>	62
16	<i>Code</i> untuk proses pelatihan dan pengujian model	63
17	<i>Code</i> untuk 10-fold cross validation	64
18	<i>Code</i> untuk proses <i>hyperparameter tuning</i> menggunakan <i>Bayesian optimization</i>	65
19	Hasil percobaan 20 <i>epoch</i> dan 32 <i>batch size</i>	66
20	Hasil percobaan 30 <i>epoch</i> dan 32 <i>batch size</i>	67
21	Hasil percobaan 50 <i>epoch</i> dan 32 <i>batch size</i>	68