

EVALUASI PENAMBAHAN *REGRESSOR* SENTIMEN ISU DPR DI PLATFORM X BERBASIS LEKSIKON DAN LLM PADA MODEL PROPHET DALAM MEMPREDIKSI IHSG

MUHAMMAD HAIKAL RASYADAN



**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**

@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Perpustakaan 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.



@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.



PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Evaluasi Penambahan *Regressor* Sentimen Isu DPR di Platform X Berbasis Leksikon dan LLM pada Model Prophet dalam Memprediksi IHSG” 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 2026

Muhammad Haikal Rasyadan
G1401221026



@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.



ABSTRAK

MUHAMMAD HAIKAL RASYADAN. Evaluasi Penambahan *Regressor* Sentimen Isu DPR di Platform X Berbasis Leksikon dan LLM pada Model Prophet dalam Memprediksi IHSG. Dibimbing oleh CICI SUHAENI dan LAILY NISSA ATUL MUALIFAH.

Perkembangan politik yang melibatkan DPR telah dikaitkan dengan fluktuasi pasar saham, menyoroti potensi nilai sentimen publik sebagai *regressor* untuk prediksi. Penelitian ini bertujuan untuk membandingkan analisis sentimen berbasis leksikon menggunakan *Indonesian sentiment* (InSet) dan pendekatan *large language model* (LLM), menganalisis tren sentimen dalam diskusi terkait DPR di platform X, dan mengevaluasi kinerja model prophet dengan penambahan *external regressor* sentimen dalam memprediksi harga penutupan indeks harga saham gabungan (IHSG). Sebanyak 45.863 *tweet* relevan dan 125 observasi harian IHSG dikumpulkan antara tanggal 2 Juni dan 28 November 2025. Analisis sentimen diawali dengan pembentukan label acuan dasar (*ground truth*) terhadap 6107 data sampel. Hasil evaluasi pada data sampel menunjukkan bahwa metode LLM (Gemini 2.5 Flash) lebih baik dibandingkan metode InSet, dengan mencapai *balanced accuracy* sebesar 0,787 dibandingkan dengan 0,405 untuk InSet. Selain itu, visualisasi *t-distributed stochastic neighbor embedding* (t-SNE) dari Word2Vec *embedding* menunjukkan konsistensi prediksi pada pendekatan LLM. Hasil analisis sentimen menunjukkan bahwa opini publik mengenai DPR didominasi oleh sentimen negatif (67,09%), yang sebagian besar mencerminkan kekhawatiran terkait isu-isu hukum dan kinerja institusi. Integrasi rasio sentimen negatif sebagai *external regressor* meningkatkan performa prediksi model prophet, yang ditunjukkan oleh penurunan nilai *mean absolute percentage error* (MAPE) pada data uji dari 3,271% menjadi 1,682%.

Kata kunci: DPR, IHSG, inset, LLM, prophet



ABSTRACT

MUHAMMAD HAIKAL RASYADAN. Evaluation of Adding DPR Issue Sentiment Regressors on the X Using Lexicon and LLM Approaches in the Prophet Model for Predicting the IHSG. Supervised by CICI SUHAENI and LAILY NISSA ATUL MUALIFAH.

Political developments involving DPR have been associated with fluctuations in the stock market, highlighting the potential value of public sentiment as an input for forecasting. This study aims to compare lexicon based sentiment analysis using Indonesian sentiment (InSet) and large language model (LLM) approaches, analyze sentiment trends in DPR related discussions on platform X, and evaluate performance of prophet model with additional of sentiment external regressor in forecasting the Indonesia composite index (IHSG) closing prices. A total of 45.863 relevant tweets and 125 daily IHSG observations were collected between June 2 and November 28, 2025. Sentimen analysis began with the creation of ground truth labels for 6107 sampled data. Evaluation results on samples data indicated LLM approach (Gemini 2.5 Flash) outperformed InSet, achieving balanced accuracy of 0,787 compared with 0,405 for InSet. Furthermore, t-distributed stochastic neighbor embedding (t-SNE) visualization of Word2Vec embeddings demonstrated labeling consistency in the LLM approach. Sentiment analysis results showed that public discourse regarding the DPR dominant negative (67,09%), largely reflecting concerns related to legal issues and institutional performance. Incorporating negative sentiment ratio as an external regressor improved prophet model's forecasting performance, reducing the mean absolute percentage error (MAPE) on the test dataset from 3,271% to 1,682%.

Keywords: DPR, IHSG, inset, LLM, prophet.



@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.

© Hak Cipta milik IPB, tahun 2026¹
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.



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

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.

Perpustakaan IPB University

EVALUASI PENAMBAHAN *REGRESSOR* SENTIMEN ISU DPR DI PLATFORM X BERBASIS LEKSIKON DAN LLM PADA MODEL PROPHET DALAM MEMPREDIKSI IHSG

MUHAMMAD HAIKAL RASYADAN

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI SARJANA STATISIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@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.

Tim Penguji pada Ujian Skripsi:

1. Eka Putri Nur Utami, S.Stat., M.Si.



Judul Skripsi : Evaluasi Penambahan *Regressor* Sentimen Isu DPR di Platform X Berbasis Leksikon dan LLM pada Model Prophet dalam Memprediksi IHSG

Nama : Muhammad Haikal Rasyadan

NIM : G1401221026

Disetujui oleh

Pembimbing 1:

Cici Suhaeni, S.Si., M.Si., Ph.D.

Pembimbing 2:

Laily Nissa Atul Kualifah, S.Si., M.Si.

Diketahui oleh

Ketua Program Studi:

Dr. Bagus Sartono, S.Si., M.Si.

NIP 197804112005011002

Tanggal Ujian:
17 Juni 2026

Tanggal Lulus:



@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.

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 November 2025 sampai bulan Juni 2026 ini ialah analisis sentimen publik dan prediksi data saham, dengan judul “Evaluasi Penambahan Regressor Sentimen Isu DPR di Platform X Berbasis Leksikon dan LLM pada Model Prophet dalam Memprediksi IHSG”.

Penyelesaian karya ilmiah ini tidak akan terwujud tanpa bantuan, dukungan, serta kontribusi berharga dari berbagai pihak yang telah berperan penting selama proses penelitian berlangsung. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Orang tua tercinta penulis, Bapak Irwan Gunanto dan Ibu Linda Puspita, serta saudara-saudara dan seluruh keluarga besar penulis. Dukungan yang tiada henti, doa yang tulus, dan kasih sayang yang tanpa syarat senantiasa menjadi sumber motivasi dan kekuatan sepanjang perjalanan ini.
2. Dosen pembimbing penulis, Cici Suhaeni, S.Si., M.Si., Ph.D., dan Laily Nissa Atul Mualifah, S.Si., M.Si., atas bimbingan yang luar biasa, arahan yang penuh wawasan, serta dukungan yang tidak pernah surut dalam penyusunan dan pelaksanaan penelitian ini.
3. Dosen penguji, Eka Putri Nur Utami, S.Stat., M.Si., atas masukan yang berharga, kritik yang membangun, dan saran-saran yang sangat bermanfaat sehingga penelitian ini dapat terus diperbaiki dan disempurnakan.
4. Moderator kolokium dan seminar, Dr. Yenni Angraini, S.Si., M.Si. dan Dr. Utami Dyah Syafitri, S.Si., M.Si., atas umpan balik konstruktif yang mendorong penulis untuk mempertajam fokus penelitian.
5. Seluruh dosen dan tenaga kependidikan Program Studi Sarjana Statistika dan Sains Data. Dedikasi mereka dalam berbagi ilmu pengetahuan serta bantuan yang diberikan selama perjalanan akademik dan proses penyusunan karya ilmiah ini sangat penulis hargai.
6. Sahabat-sahabat terdekat penulis, khususnya Rakha, Alista, Ghonniyu, Azanti, Bintang, Bulan, Yumna, Shyfa, Intan, Jodi, Oci, Rakesh, dan Laras, atas dukungan, kebersamaan, serta berbagai diskusi yang penuh semangat telah menemani dan memberikan warna dalam perjalanan ini.
7. Supraption Fams, Tor Monitor, Qluv Malam, Majelis Taklim KDW yang telah memberikan dukungan, doa, kebersamaan, serta semangat kepada penulis.
8. Kucing kesayangan penulis, cibul, kibu, ciko, dan cipung, yang telah menjadi sumber dukungan emosional selama proses penyusunan karya ilmiah
9. Marinestic Statistika 59, GSB, Departemen *Survey and Research*, dan seluruh pihak yang telah memberikan pengalaman, dan dukungan yang berharga.

Akhir kata, penulis berharap karya ilmiah ini dapat memberikan manfaat bagi pihak-pihak yang membutuhkan serta berkontribusi terhadap pengembangan ilmu pengetahuan.

Bogor, Agustus 2026

Muhammad Haikal Rasyadan



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

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.

Perpustakaan IPB University

DAFTAR ISI

DAFTAR TABEL	xiv
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xv
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	4
1.4 Manfaat	4
II TINJAUAN PUSTAKA	5
2.1 Indeks Harga Saham Gabungan (IHSG)	5
2.2 Leksikon <i>Indonesian Sentiment</i> (InSet)	5
2.3 <i>Large Language Model</i> (LLM)	6
2.4 <i>Word Embedding</i> Word2Vec	9
2.5 <i>t-Distributed Stochastic Neighbor Embedding</i> (t-SNE)	10
2.6 Model Prophet	11
III METODE	15
3.1 Data	15
3.2 Prosedur Analisis Data	15
IV HASIL DAN PEMBAHASAN	23
4.1 Pengumpulan dan <i>Screening</i> Data <i>Tweet</i>	23
4.2 <i>Preprocessing</i> Data <i>Tweet</i>	24
4.3 Eksplorasi Data	25
4.4 Pembentukan Label <i>Ground Truth</i>	28
4.5 Prediksi Sentimen	31
4.6 Evaluasi Metode Analisis Sentimen	33
4.7 Implementasi dan Evaluasi Metode Analisis Sentimen Terbaik	37
4.8 <i>Feature Engineering</i> Sentimen	40
4.9 Identifikasi <i>Delay Effect</i>	42
4.10 Prediksi Harga Penutupan IHSG dengan Model Prophet	44
4.11 Evaluasi Komparatif Model Prophet	49
V SIMPULAN DAN SARAN	51
5.1 Simpulan	51
5.2 Saran	51
DAFTAR PUSTAKA	53
LAMPIRAN	57
RIWAYAT HIDUP	73

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.



DAFTAR TABEL

1	Contoh daftar kata pada leksikon InSet (Musfiroh <i>et al.</i> 2021)	6
2	Interpretasi hasil pengukuran kappa (Landis dan Koch 1977)	17
3	Contoh atribut data <i>tweet</i> hasil <i>web crawling</i>	23
4	Contoh data <i>tweet</i> yang tidak relevan dan direduksi	24
5	Contoh data tahapan <i>preprocessing</i>	25
6	Contoh penentuan label akhir anotasi manual	29
7	Cohen's Kappa	29
8	Fleiss' Kappa	30
9	Contoh hasil prediksi sentimen menggunakan leksikon InSet	31
10	Contoh hasil prediksi sentimen menggunakan LLM Gemini 2.5 <i>Flash</i>	33
11	Evaluasi metode analisis sentimen	37
12	Pengecekan awal stasioneritas ragam dan rata-rata	42
13	Hasil penanganan ketidakstasioneran data	43
14	Evaluasi model prophet	50
15	Uji Diebold-Mariano	50

DAFTAR GAMBAR

1	Mekanisme <i>self attention</i> (Vaswani <i>et al.</i> 2027)	9
2	<i>Workflow</i> LLM dengan platform n8n	18
3	Tren jumlah <i>tweet</i> relevan	26
4	Tren harga penutupan IHSG	27
5	<i>Wordcloud</i> hasil <i>text preprocessing</i>	27
6	Proporsi kategori sentimen hasil pelabelan manual	28
7	Proporsi kategori sentimen hasil akhir pelabelan manual	30
8	Proporsi kategori sentimen hasil prediksi leksikon InSet	31
9	Proporsi kategori sentimen hasil prediksi LLM	32
10	Perbandingan hasil pelabelan sentimen	34
11	Perbandingan <i>wordcloud</i> analisis sentimen. (A) <i>Ground truth</i> negatif; (B) InSet negatif; (C) LLM negatif; (D) <i>Ground truth</i> netral; (E) InSet netral; (F) LLM netral; (G) <i>Ground truth</i> positif; (H) InSet positif; (I) LLM positif	35
12	Perbandingan matriks konfusi analisis sentimen. (A) <i>ground truth</i> dan leksikon InSet; (B) <i>ground truth</i> dan LLM	36
13	Proporsi kategori sentimen pada data nonsampel	37
14	Perbandingan <i>wordcloud</i> pelabelan sentimen dengan LLM. (A) Sentimen negatif data dalam sampel; (B) Sentimen negatif data luar sampel; (C) Sentimen netral data dalam sampel; (D) Sentimen netral data luar sampel; (E) Sentimen positif data dalam sampel; (F) Sentimen positif data luar sampel	38
15	Visualisasi t-SNE sebaran kata sentimen	40
16	Tren rasio sentimen negatif	41
17	Tren rasio sentimen positif	41



18	Plot CCF (A) rasio sentimen negatif, (B) rasio sentimen positif	43
19	Grafik perbandingan (A) harga penutupan IHSG, (B) rasio sentimen negatif <i>lag</i> ke-3	44
20	Grafik pembandingan data latih dan data uji	45
21	Hasil prediksi model prophet tanpa <i>external regressor</i>	46
22	Hasil prediksi model prophet dengan <i>external regressor</i>	48

DAFTAR LAMPIRAN

1	<i>Prompt LLM</i> pada n8n	57
2	Hasil <i>hyperparameter tuning</i> pada model prophet tanpa <i>external regressor</i>	58
3	Hasil <i>hyperparameter tuning</i> pada model prophet dengan <i>external regressor</i>	64
4	Persamaan penuh model prophet tanpa <i>external regressor</i>	70
5	Persamaan penuh model prophet dengan <i>external regressor</i>	71