

**PREDIKSI PERGERAKAN HARGA SAHAM
BERDASARKAN LAPORAN TAHUNAN PERUSAHAAN TERBUKA
MENGUNAKAN METODE *CONTEXTUAL EMBEDDING BERT*
DAN *LONG SHORT-TERM MEMORY***

SKRIPSI

diajukan untuk memenuhi sebagian syarat
untuk memperoleh gelar Sarjana Komputer
Program Studi Ilmu Komputer



Oleh
Yosafat
NIM 2009929

**PROGRAM STUDI ILMU KOMPUTER
FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS PENDIDIKAN INDONESIA
2024**

**PREDIKSI PERGERAKAN HARGA SAHAM
BERDASARKAN LAPORAN TAHUNAN PERUSAHAAN TERBUKA
MENGUNAKAN METODE *CONTEXTUAL EMBEDDING BERT*
DAN *LONG SHORT-TERM MEMORY***

Oleh
Yosafat
NIM 2009929

Sebuah skripsi yang diajukan untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Komputer pada Fakultas Pendidikan Matematika dan Ilmu Pengetahuan
Alam

© Yosafat 2024
Universitas Pendidikan Indonesia
Agustus 2024

Hak Cipta dilindungi undang-undang.
Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian,
dengan dicetak ulang, difoto kopi, atau cara lainnya tanpa ijin dari penulis.

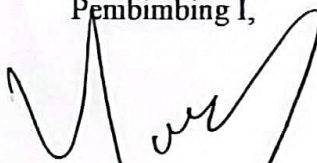
YOSAFAT

2009929

**PREDIKSI PERGERAKAN HARGA SAHAM
BERDASARKAN LAPORAN TAHUNAN PERUSAHAAN TERBUKA
MENGUNAKAN METODE *CONTEXTUAL EMBEDDING* BERT DAN
*LONG SHORT-TERM MEMORY***

DISETUJUI DAN DISAHKAN OLEH PEMBIMBING:

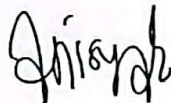
Pembimbing I,



Dr. Yudi Wibisono, M.T.

NIP: 197507072003121003

Pembimbing II,



Ani Anisvah, M.T.

NIP: 920200419930811201

Mengetahui,

Ketua Program Studi Ilmu Komputer



Dr. Muhamad Nursalman, M. T.

NIP: 197909292006041002

PERNYATAAN

Saya menyatakan bahwa skripsi yang berjudul “**Prediksi Pergerakan Harga Saham Berdasarkan Laporan Tahunan Perusahaan Terbuka Menggunakan Metode *Contextual Embedding BERT* dan *Long Short-Term Memory***” ini sepenuhnya merupakan hasil karya saya sendiri. Tidak ada bagian di dalamnya yang merupakan plagiat dari karya orang lain dan saya tidak melakukan penyalinan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini, saya siap menanggung risiko/sanksi yang dijatuhkan kepada saya apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya ini atau ada klaim dari pihak lain terhadap keaslian karya saya ini.

Bandung, Agustus 2024

Yang membuat pernyataan



Yosafat

2009929

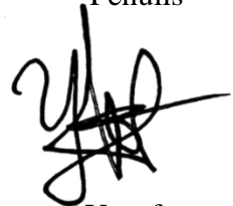
KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas segala limpahan kasih, karunia, dan kehendak-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “**Prediksi Pergerakan Harga Saham Berdasarkan Laporan Tahunan Perusahaan Terbuka Menggunakan Metode *Contextual Embedding BERT* dan *Long Short-Term Memory***” ini dengan baik. Penulis menyusun skripsi ini ditujukan untuk memenuhi dan melengkapi salah satu syarat untuk mendapatkan gelar Sarjana Komputer pada Program Studi Ilmu Komputer, Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam Universitas Pendidikan Indonesia (FPMIPA UPI).

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat banyak kekurangan dan keterbatasan yang perlu disempurnakan. Oleh karena itu, penulis sangat mengharapkan saran maupun kritik yang membangun dari pembaca. Semoga skripsi ini bermanfaat bagi masyarakat pada umumnya dan bagi akademiperkembangan ilmu pengetahuan.

Bandung, Agustus 2024

Penulis



Yosafat

2009929

UCAPAN TERIMA KASIH

Dalam proses penyelesaian skripsi ini, penulis tentunya tidak lepas dari bantuan dan dorongan dari pihak-pihak yang telah membantu, baik secara langsung maupun tidak langsung. Oleh karena itu, penulis pada kesempatan ini mengucapkan terima kasih kepada:

1. Ibu, kakak, dan adik tercinta yang senantiasa mendoakan, memberikan dukungan materil dan non-materil, dorongan, semangat, dan motivasi selama masa perkuliahan dan masa penyelesaian skripsi ini.
2. Seluruh anggota keluarga besar, terutama om, tante, kakek, dan nenek yang senantiasa mendoakan, memberikan dukungan dan motivasi selama masa perkuliahan.
3. Bapak Dr. Yudi Wibisono, MT., selaku pembimbing I yang telah bersedia memberikan masukan, saran, dan kritik yang membangun serta meluangkan waktu, tenaga, dan pikiran dalam penyelesaian skripsi ini.
4. Ibu Ani Anisyah, M.T., selaku pembimbing II yang telah memberikan masukan dan saran dalam penyelesaian skripsi ini.
5. Bapak Eddy Prasetyo Nugroho, M.T., selaku dosen pembimbing akademik yang telah memberikan bimbingan selama masa perkuliahan.
6. Bapak Dr. Muhamad Nursalman, M.T., selaku Ketua Program Studi Ilmu Komputer FPMIPA Universitas Pendidikan Indonesia.
7. Seluruh dosen Departemen Pendidikan Ilmu Komputer yang telah membimbing dan memberikan ilmu yang bermanfaat pada penulis selama masa perkuliahan.
8. Teman-teman seperjuangan Program Studi Ilmu Komputer Angkatan 2020 yang telah mendukung penulis selama masa perkuliahan.

Penulis merasa bahwa ungkapan terima kasih tidak cukup untuk menggambarkan rasa syukur atas segala bantuan yang telah diterima. Semoga Tuhan Yang Maha Esa membalas segala kebaikan dan ketulusan yang telah diberikan kepada penulis dalam proses penyelesaian skripsi ini.

**PREDIKSI PERGERAKAN HARGA SAHAM
BERDASARKAN LAPORAN TAHUNAN PERUSAHAAN TERBUKA
MENGUNAKAN METODE *CONTEXTUAL EMBEDDING* BERT
DAN *LONG SHORT-TERM MEMORY***

ABSTRAK

Pasar saham bersifat kompleks dan cenderung memiliki volatilitas yang tinggi. Hal ini menyebabkan proses pengambilan keputusan investasi menjadi sulit. Informasi tekstual pada laporan tahunan perusahaan terbuka dapat dijadikan data untuk memprediksi pergerakan harga saham. Namun, prediksi ini memiliki tantangan tersendiri karena perlu memproses data teks yang memiliki sifat tidak terstruktur. Penelitian ini mengusulkan metode *contextual embedding* BERT untuk mengubah teks laporan tahunan menjadi representasi vektor dan melakukan klasifikasi menggunakan *Long Short-Term Memory* (LSTM). *Dataset* yang digunakan terdiri dari 304 laporan tahunan dari 34 perusahaan terbuka yang dikumpulkan mulai dari tahun 2005 sampai dengan tahun 2020. Berdasarkan eksperimen yang dilakukan terhadap 4 skenario: (1) penggunaan data; (2) *baseline*; (3) pemotongan teks (*truncation*); dan (4) *hyperparameter*, diperoleh temuan yang menunjukkan bahwa laporan tahunan yang memiliki pengaruh signifikan terhadap pergerakan harga saham adalah laporan tahunan yang terbit 2-3 tahun sebelumnya. Selain itu, bagian akhir (*tail*) pada teks laporan tahunan mengandung informasi penting dan berpengaruh signifikan terhadap pergerakan harga saham. Model klasifikasi terbaik menghasilkan F1-Score untuk kelas harga saham naik yaitu 0.78 dan menggunakan data 3 tahun setelah laporan tahunan diterbitkan, menggunakan bagian akhir (*tail*) teks, dan *batch size* 16. Sedangkan untuk kelas harga saham turun, model terbaik menghasilkan F1-Score yaitu 0.57 dan menggunakan data 2 tahun setelah laporan tahunan terbit, menggunakan bagian akhir (*tail*) teks, dan *batch size* 16.

Kata Kunci: Prediksi Pergerakan Harga Saham, Laporan Tahunan, *Contextual Embedding*, BERT, LSTM

**PREDIKSI PERGERAKAN HARGA SAHAM
BERDASARKAN LAPORAN TAHUNAN PERUSAHAAN TERBUKA
MENGUNAKAN METODE *CONTEXTUAL EMBEDDING BERT*
DAN *LONG SHORT-TERM MEMORY***

ABSTRACT

The stock market is complex and tends to have high volatility. This makes the investment decision-making process difficult. Textual information in the annual reports of listed companies can be used as data to predict stock price movements. However, this prediction has its own challenges because it needs to process text data that has an unstructured nature. This research proposes a contextual embedding BERT method to convert annual report text into vector representation and perform classification using Long Short-Term Memory (LSTM). The dataset used consists of 304 annual reports from 34 public companies collected from 2005 to 2020. Based on experiments conducted on 4 scenarios: (1) data usage; (2) baseline; (3) truncation; and (4) hyperparameter, the findings show that annual reports that have a significant influence on stock price movements are annual reports published 2-3 years earlier. In addition, the tail of the annual report text contains important information and has a significant effect on stock price movements. The best classification model produces F1-Score for the rising price class which is 0.78 and uses data 3 years after the annual report is published, uses the tail part of the text, and batch size 16. As for the descending class, the best model produces F1-Score which is 0.57 and uses data 2 years after the annual report is published, uses the tail part of the text, and batch size 16.

Keywords: Stock Price Movement Prediction, Annual Report, Contextual Embedding, BERT, LSTM

DAFTAR ISI

PERNYATAAN	i
KATA PENGANTAR	ii
UCAPAN TERIMA KASIH	iii
ABSTRAK	iv
ABSTRACT	v
DAFTAR ISI	vi
DAFTAR TABEL	viii
DAFTAR GAMBAR	ix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	7
1.3 Tujuan Penelitian.....	7
1.4 Manfaat Penelitian.....	7
1.5 Batasan Penelitian	8
1.6 Sistematika Penulisan.....	8
BAB II KAJIAN PUSTAKA	10
2.1 Penelitian Terkait	10
2.2 Pasar Saham	12
2.3 Saham	13
2.4 Perusahaan Terbuka (Tbk)	18
2.5 Laporan Tahunan.....	20
2.6 <i>Natural Language Processing (NLP)</i>	24
2.7 <i>Deep Learning</i>	26
2.8 <i>Transformers</i>	28
2.9 <i>Word Embedding</i>	31
2.10 <i>Contextual Embedding</i>	35

2.11 <i>Bidirectional Encoder Representations from Transformers (BERT)</i>	38
2.12 <i>Recurrent Neural Network (RNN)</i>	41
2.13 <i>Long Short-Term Memory (LSTM)</i>	44
2.13.1 <i>Forget Gate</i>	47
2.13.2 <i>Input Gate</i>	48
2.13.3 <i>Cell State</i>	49
2.13.4 <i>Output Gate</i>	49
2.14 <i>Baseline</i>	50
2.15 <i>Hyperparameter</i>	51
2.16 <i>Evaluasi Model</i>	52
BAB III METODE PENELITIAN	55
3.1 <i>Desain Penelitian</i>	55
3.2 <i>Alat dan Bahan Penelitian</i>	59
3.2.1 <i>Alat Penelitian</i>	60
3.2.2 <i>Bahan Penelitian</i>	60
BAB IV HASIL DAN PEMBAHASAN	63
4.1 <i>Hasil</i>	63
4.1.1 <i>Pembuatan Dataset</i>	63
4.1.2 <i>Persiapan Dataset</i>	69
4.1.3 <i>Pembangunan Model Klasifikasi</i>	75
4.1.4 <i>Skenario Eksperimen</i>	76
4.1.5 <i>Eksperimen</i>	79
4.1.6 <i>Pengembangan Aplikasi</i>	87
4.2 <i>Pembahasan</i>	90
BAB V KESIMPULAN DAN SARAN	97
5.1 <i>Kesimpulan</i>	97
5.2 <i>Saran</i>	98
DAFTAR PUSTAKA	99
LAMPIRAN	106

DAFTAR TABEL

Tabel 2. 1 <i>Confusion Matrix</i>	53
Tabel 2. 2 Deskripsi Metrik Evaluasi <i>Confusion Matrix</i>	53
Tabel 3. 1 Daftar Laporan Tahunan Emiten.....	61
Tabel 4. 1 Periode Pengumpulan Data Laporan Tahunan.....	63
Tabel 4. 2 Contoh Data Harga Saham.....	66
Tabel 4. 3 Batas Penentuan Label	68
Tabel 4. 4 Contoh Sebagian <i>Dataset</i> Setelah Pelabelan Data untuk Data Y+1	68
Tabel 4. 5 Contoh Hasil Praproses Teks	69
Tabel 4. 6 Contoh Hasil Pemotongan Teks (<i>Truncation</i>)	71
Tabel 4. 7 Contoh Hasil Tokenisasi	72
Tabel 4. 8 Contoh Hasil Vektorisasi	73
Tabel 4. 9 Pembagian Data	75
Tabel 4. 10 Informasi <i>Dataset</i>	77
Tabel 4. 11 Hasil Eksperimen <i>Baseline</i> Naïve Bayes dan TF-IDF <i>Embedding</i>	79
Tabel 4. 12 Hasil Eksperimen <i>Baseline</i> Naïve Bayes dan Word2Vec <i>Embedding</i>	81
Tabel 4. 13 Hasil Eksperimen <i>Baseline</i> Naïve Bayes dan <i>Contextual Embedding</i> BERT.....	82
Tabel 4. 14 Hasil Eksperimen <i>Hyperparameter</i> LSTM.....	83
Tabel 4. 15 Hasil Model <i>Baseline</i>	91
Tabel 4. 16 Hasil Model LSTM Terbaik untuk Kelas Harga Saham Naik dan Kelas Harga Saham Turun	92

DAFTAR GAMBAR

Gambar 2. 1 Grafik Harga Saham BBCA Tahun 2004-2024	15
Gambar 2. 2 Skenario Pelabelan <i>Volatility Based Threshold</i>	18
Gambar 2. 3 Sampul Laporan Tahunan PT Bank Central Asia Tbk Tahun 2023.	22
Gambar 2. 4 Tata Kelola Perusahaan PT Bank Central Asia Tbk Tahun 2023	23
Gambar 2. 5 Komponen NLP.....	25
Gambar 2. 6 Arsitektur <i>Transformers</i> (dimodifikasi)	29
Gambar 2. 7 <i>Skip-Gram</i>	32
Gambar 2. 8 CBOW	33
Gambar 2. 9 Matriks Kejadian Bersama GloVe	34
Gambar 2. 10 <i>Masked Language Model</i> (MLM)	40
Gambar 2. 11 <i>Next Sentence Prediction</i> (NSP).....	41
Gambar 2. 12 Arsitektur RNN	42
Gambar 2. 13 Arsitektur LSTM	45
Gambar 2. 14 Fungsi Sigmoid.....	46
Gambar 2. 15 Fungsi Tanh	46
Gambar 2. 16 <i>Forget Gate</i>	47
Gambar 2. 17 <i>Input Gate</i>	48
Gambar 2. 18 <i>Cell State</i>	49
Gambar 2. 19 <i>Output Gate</i>	49
Gambar 2. 20 Perhitungan Sederhana LSTM untuk Prediksi Harga Saham	50
Gambar 3. 1 Desain Penelitian	55
Gambar 4. 1 Halaman Daftar Laporan Tahunan untuk Kode Emiten BBCA.....	65
Gambar 4. 2 Hasil Pengumpulan Data Laporan Tahunan Emiten	65
Gambar 4. 3 Diagram Alur Pelabelan Data.....	67
Gambar 4. 4 Contoh Hasil <i>Encoding</i>	73
Gambar 4. 5 Alur Proses Pelatihan dan Pengujian.....	75
Gambar 4. 6 Diagram Arsitektur <i>Long Short-Term Memory</i> (LSTM)	76
Gambar 4. 7 Halaman Antarmuka WebApp Prediksi Pergerakan Harga Saham Menggunakan Laporan Tahunan.....	87

Gambar 4. 8 Antarmuka Fungsional Meng- <i>upload</i> Laporan Tahunan	88
Gambar 4. 9 Antarmuka Fungsional Memilih Periode Waktu Prediksi	89
Gambar 4. 10 Antarmuka Fungsional Prediksi dengan Hasil Prediksi Harga Saham Turun	89
Gambar 4. 11 Antarmuka Fungsional Prediksi dengan Hasil Prediksi Harga Saham Tetap.....	90
Gambar 4. 12 Antarmuka Fungsional Prediksi dengan Hasil Prediksi Harga Saham Naik	90
Gambar 4. 13 Contoh Bagian Akhir (<i>Tail</i>) Laporan Tahunan Perusahaan Terbuka Berupa Informasi Catatan Keuangan Perusahaan	94
Gambar 4. 14 Contoh Bagian Akhir (<i>Tail</i>) Laporan Tahunan Perusahaan Terbuka Berupa Informasi Produk/Layanan Perusahaan	94
Gambar 4. 15 <i>Training</i> dan <i>Test Loss History</i> Model LSTM Terbaik untuk Kelas Harga Saham Naik	95
Gambar 4. 16 Grafik <i>Training</i> dan <i>Test Loss History</i> Model LSTM Terbaik untuk Kelas Harga Saham Naik	95
Gambar 4. 17 <i>Classification Report</i> Model LSTM Terbaik untuk Kelas Harga Saham Naik	96
Gambar 4. 18 <i>Confusion Matrix</i> Model LSTM Terbaik untuk Kelas Harga Saham Naik	96

DAFTAR PUSTAKA

- Ameisen, E. (2018). Always start with a stupid model, no exceptions. Diambil 15 Mei 2024, dari Medium website: <https://blog.insightdatascience.com/always-start-with-a-stupid-model-no-exceptions-3a22314b9aaa>
- Baker, S. R., Bloom, N., Davis, S. J., Sammon, M., Wang, S., Lee, A., ... Stanford University, N. (2022). *What Triggers Stock Market Jumps?* Diambil dari <https://stockmarketjumps.com/research/>.
- Balbi, S., Stawinoga, A., & Triunfo, N. (2012). *Text Mining tools for extracting knowledge from Firms Annual Reports*. Diambil dari www.blue-ets.istat.it
- Bhatia, T., Colianni, S., Korf, T., & Katz, E. (2019). A Dynamic Labeling Approach for Financial Assets Forecasting — Neuravest. Diambil 30 Juli 2024, dari <https://www.neuravest.net/a-dynamic-labeling-approach-for-financial-assets-forecasting-2/>
- Chen, Q. (2021). *Stock Movement Prediction with Financial News using Contextualized Embedding from BERT*.
- Dash, N. S. (2008). Context and Contextual Word Meaning Niladri Sekhar Dash. *SKASE Journal of Theoretical Linguistics*, 5(2).
- Devlin, J., Chang, M.-W., Lee, K., Google, K. T., & Language, A. I. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. Diambil dari <https://github.com/tensorflow/tensor2tensor>
- Dirdjosisworo, S. (1997). *Hukum Perusahaan Mengenai Bentuk-bentuk Perusahaan (badan usaha) di Indonesia*. Mandar Maju.
- Doucette, J. A., & Cohen, R. (2015). *Content of Annual Reports as a Predictor for Long Term Stock Price Movements*. Diambil dari www.aaai.org

- Fazlija, B., & Harder, P. (2022). Using Financial News Sentiment for Stock Price Direction Prediction. *Mathematics* 2022, Vol. 10, Page 2156, 10(13), 2156. <https://doi.org/10.3390/MATH10132156>
- Glend, V. (2022, Juli 14). Annual Report: Pengertian, Fungsi, Komponen hingga Syaratnya. Diambil 19 Mei 2023, dari <https://www.hashmicro.com/id/blog/annual-report/>
- Hájek, P., Olej, V., & Myšková, R. (2013). *Forecasting Stock Prices using Sentiment Information in Annual Reports-A Neural Network and Support Vector Regression Approach*.
- Handini, S., & Astawinetu, E. D. (2020). *Teori Portofolio dan Pasar Modal Indonesia*. Scopindo Media Pustaka.
- Hermuningsih, S. (2012). *Pengantar Pasar Modal Indonesia*. Yogyakarta: UPP STIM YKPN.
- Heydarian, M., Doyle, T. E., & Samavi, R. (2022). MLCM: Multi-Label Confusion Matrix. *IEEE Access*, 10, 19083–19095. <https://doi.org/10.1109/ACCESS.2022.3151048>
- Hidayat, T. (2010). *Buku Pintar Investasi*. Jakarta: Mediakita.
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/NECO.1997.9.8.1735>
- Hutter, F., Hoos, H., & Leyton-Brown, K. (2014). *An Efficient Approach for Assessing Hyperparameter Importance*.
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685–695. <https://doi.org/10.1007/S12525-021-00475-2/TABLES/2>

- Jeany, G., & Tjun, L. T. (2016). Pengaruh Current Ratio (CR), Earnings per Share (EPS) dan Price Earnings Ratio Terhadap Harga Saham: Studi pada Perusahaan yang Terdaftar di Bursa Efek Indonesia Periode 2012-2014. *Jurnal Akuntansi*, 8(1), 131–156. Diambil dari www.kommpasiana.com
- Jogiyanto, H. (2010). *Teori Portofolio dan Analisis Investasi*. Yogyakarta: BPF.
- Joshua, Z.-G. H., Huang, X., Hao Mou, P., Duan Li, P., Qi Wu, P., & Yabo Xu, P. (2022). *BERT-based Financial Sentiment Index and LSTM-based Stock Return Predictability*. Diambil dari www.datastory.com.cn
- Jumingan. (2011). *Analisis Laporan Keuangan*. Jakarta: PT Bumi Aksara.
- Kartika, A., Kendeng, J. V., & Ngisor, B. (2010). *Volatilitas Harga Saham di Indonesia dan Malaysia*. 12(1), 17–26.
- Kasmir. (2010). *Pengantar Manajemen Keuangan Edisi Pertama* (Ed. I). Jakarta: Prenada Media Group.
- Khurana, D., Koli, A., Khatker, K., & Singh, S. (2023). Natural language processing: state of the art, current trends and challenges. *Multimedia Tools and Applications*, 82(3), 3713–3744. <https://doi.org/10.1007/S11042-022-13428-4>
- Kilimci, Z. H., Uysal, M., & Othan, D. (2019). *Financial Sentiment Analysis for Predicting Direction of Stocks using Bidirectional Encoder Representations from Transformers (BERT) and Deep Learning Models*. <https://doi.org/10.17758/URUAE8.UL12191013>
- Kohar, A., Ahmar, N., & Suratno. (2019). SENSITIVITAS FAKTOR EKONOMI MAKRO DAN MIKRO DALAM MEMPREDIKSI VOLATILITAS HARGA SAHAM PERUSAHAAN SEKTOR INDUSTRI FOOD & BEVERAGES. *JIAFE (Jurnal Ilmiah Akuntansi Fakultas Ekonomi)*, 4(1), 85–100. <https://doi.org/10.34204/JIAFE.V4I1.1080>

- Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). *IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP*. Diambil dari <https://huggingface.co/>
- Krenker, A., Bešter, J., & Kos, A. (2011). Introduction to the artificial neural networks. *Artificial Neural Networks: Methodological Advances and Biomedical Applications. InTech*, 1–18.
- Lecun, Y., Bengio, Y., Hinton, G. ;, Gregor K And Lecun, Y., Icml ; F, K. D., Philbin, J., ... Chen, Z. (2018). PERSPECTIVES Special Topic: Machine Learning Deep learning for natural language processing: advantages and challenges. *National Science Review*, 5(1), 22–24. <https://doi.org/10.1093/nsr/nwx099>
- Liddy, E. D. (2001). *Natural Language Processing*. Diambil dari <https://surface.syr.edu/istpub>
- Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance*, 66(1), 35–65. <https://doi.org/10.1111/J.1540-6261.2010.01625.X>
- Maxime. (2019, Januari 5). What is a Transformer?. An Introduction to Transformers and... | by Maxime | Inside Machine learning | Medium. Diambil 10 Mei 2024, dari Medium website: <https://medium.com/inside-machine-learning/what-is-a-transformer-d07dd1fbec04>
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space*. Diambil dari <http://ronan.collobert.com/senna/>
- Nyuytiybiy, K. (2020). Parameters, Hyperparameters, Machine Learning | Towards Data Science. Diambil 22 Mei 2024, dari Medium website: <https://towardsdatascience.com/parameters-and-hyperparameters-aa609601a9ac>

- Parker, T. (2023, Mei 2). Cognitive vs. Emotional Investing Bias: What's the Difference? Diambil 30 April 2024, dari <https://www.investopedia.com/articles/investing/051613/behavioral-bias-cognitive-vs-emotional-bias-investing.asp>
- Pascanu, R., Gulcehre, C., Cho, K., & Bengio, Y. (2014). *How to Construct Deep Recurrent Neural Networks*.
- Pascanu, R., Mikolov, T., & Bengio, Y. (2013). *On the difficulty of training recurrent neural networks*.
- Patar, A., Darminto, & Saifi, M. (2014). *Faktor Internal dan Eksternal yang Mempengaruhi Pergerakan Harga Saham (Studi Pada Saham-Saham Indeks LQ45 Periode 2009–2013)*. Brawijaya University.
- Pennington, J., Socher, R., & Manning, C. D. (2014). *GloVe: Global Vectors for Word Representation*. Diambil dari <http://nlp>.
- Permata, C. P., & Ghoni, M. A. (2019). PERANAN PASAR MODAL DALAM PEREKONOMIAN NEGARA INDONESIA. *Jurnal AkunStie (JAS)*, 5(2), 50–61. <https://doi.org/10.32767/JAS.V5I2.680>
- Pradhana, R. W. (2018, Juli 5). Pengaruh Financial Literacy, Cognitive Bias, dan Emotional Bias terhadap Keputusan Investasi (Studi pada Investor Galeri Investasi Universitas Negeri Surabaya). Diambil 25 April 2024, dari Jurnal Ilmu Manajemen (JIM) website: <https://ejournal.unesa.ac.id/index.php/jim/article/view/23849>
- Republik Indonesia. *Undang-Undang Nomor. 40 Tahun 2007 tentang Perseroan Terbatas.*, (2007).
- Saliman, A. R. (2005). *Hukum Bisnis Untuk Perusahaan*. Kencana.
- Scott, W. R. (1997). *Financial accounting theory (Vol 3)*. NJ: Prentice Hall Upper Saddle River.

- Sjahrial, D. (2012). *Pengantar Manajemen Keuangan* (Edisi Keempat). Jakarta: Mitra Wacana Media.
- Sulindawati, N. L. G. E. (2017). *Manajemen Keuangan: Sebagai Dasar Pengambilan Keputusan Bisnis* (Edisi 1). Depok: Rajawali Pers.
- Sun, C., Qiu, X., Xu, Y., & Huang, X. (2020). *How to Fine-Tune BERT for Text Classification?*
- Sunariyah. (2006). *Pengetahuan Pasar Modal*. Yogyakarta: YKPN.
- Sunariyah. (2011). *Pengantar Pengetahuan Pasar Modal*. Yogyakarta: UPP AMP YKPN.
- Sutskever, I., Martens, J., & Hinton, G. (2011). *Generating Text with Recurrent Neural Networks*.
- Tandelilin, E. (2017). *Pasar Modal: Manajemen Portofolio dan Investasi* (Cetakan Pertama). Yogyakarta: PT Kanisius.
- TimCole. (2022, Mei 26). Differences Between Financial Statements and Annual Reports. Diambil 25 April 2024, dari <https://timcole.com.sg/differences-between-financial-statements-and-annual-reports/>
- Vaswani, A., Brain, G., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., ... Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural Information Processing Systems*, 30.
- Vijay Gaikwad, S., Chaugule, A., & Patil, P. (2014). Text Mining Methods and Techniques. *International Journal of Computer Applications*, 85(17), 975–8887.
- Wang, S., Bai, Y., Ji, T., Fu, K., Wang, L., & Lu, C.-T. (2023). *Stock Movement and Volatility Prediction from Tweets, Macroeconomic Factors and Historical Prices*. Diambil dari <https://github.com/hao1zhao/Bigdata23>.

- Wang, Y., Cui, L., & Zhang, Y. (2019). *Using Dynamic Embeddings to Improve Static Embeddings*.
- Wardani, S., & Fitriati, R. (2011). Analisis Komparasi Profitabilitas Sebelum dan Sesudah Penawaran Umum Saham Perdana. *BISNIS & BIROKRASI: Jurnal Ilmu Administrasi dan Organisasi*, 17(2), 1. <https://doi.org/10.20476/jbb.v17i2.630>
- Widoatmodjo, S. (1996). *Cara sehat investasi di pasar modal*. Jakarta: Jurnalindo Aksan Grafika.
- Wisniewski, T. P., & Yekini, L. S. (2014). Predicting Stock Market Returns Based on the Content of Annual Report Narrative: A New Anomaly. *SSRN Electronic Journal*. <https://doi.org/10.2139/SSRN.2474061>
- Zindel, M., Zindel, T., & Quirino, M. (2014). *Cognitive Bias and their Implications on the Financial Market*.