

**PREDIKSI PERGERAKAN HARGA SAHAM BERDASARKAN BERITA
MENGUNAKAN METODE *LONG SHORT-TERM MEMORY***

SKRIPSI

Diajukan untuk Memenuhi Sebagian dari
Syarat Memperoleh Gelar Sarjana Komputer
Program Studi Ilmu Komputer



Oleh

Fenny Feronika Maharani

1703015

**PROGRAM STUDI ILMU KOMPUTER
DEPARTEMEN PENDIDIKAN ILMU KOMPUTER
FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS PENDIDIKAN INDONESIA
2022**

**PREDIKSI PERGERAKAN HARGA SAHAM BERDASARKAN BERITA
MENGUNAKAN METODE *LONG SHORT-TERM MEMORY***

Oleh

Fenny Feronika Maharani

NIM 1703015

Sebuah Skripsi yang Diajukan untuk Memenuhi Salah Satu Syarat Memperoleh
Gelar Sarjana Komputer pada Fakultas Pendidikan Matematika dan Ilmu
Pengetahuan Alam

©

Universitas Pendidikan Indonesia

Desember 2021

Hak Cipta Dilindungi Undang-Undang

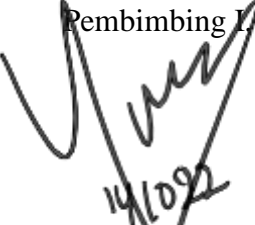
Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian, dengan dicetak
ulang, difoto kopi, atau cara lainnya tanpa izin dari penulis

FENNY FERONIKA MAHARANI


1703015

**PREDIKSI PERGERAKAN HARGA SAHAM BERDASARKAN BERITA
MENGUNAKAN METODE *LONG SHORT-TERM MEMORY***

DISETUJUI DAN DISAHKAN OLEH PEMBIMBING:

Pembimbing I

14/10/22
Dr. Yudi Wibisono, M.T.
NIP. 197507072003121003

Pembimbing II,


Herbert Siregar, M.T.
NIP. 197005022008121001

Mengetahui,

Ketua Departemen Pendidikan Ilmu Komputer



Dr. Rani Megasari, M.T.
NIP. 198705242014042002

PERNYATAAN

Saya menyatakan bahwa skripsi yang berjudul “**Prediksi Pergerakan Harga Saham Berdasarkan Berita Menggunakan Metode *Long Short-Term Memory***” ini sepenuhnya hasil karya saya sendiri. Tidak ada bagian di dalamnya yang merupakan plagiat dari karya orang lain dan saya tidak melakukan penjiplakan 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, Maret 2022

Yang membuat pernyataan

Fenny Feronika Maharani

1703015

PREDIKSI PERGERAKAN HARGA SAHAM BERDASARKAN BERITA MENGUNAKAN METODE *LONG SHORT-TERM MEMORY*

Oleh

Fenny Feronika Maharani — fenny@upi.edu

1703015

ABSTRAK

Pasar saham memiliki perilaku yang kompleks dan dinamis, sehingga pengambilan keputusan investasi di pasar saham merupakan tugas yang sulit. Informasi publik berupa berita dapat dijadikan sebagai data untuk memprediksi pergerakan harga saham, namun prediksi pergerakan harga saham berdasarkan berita memiliki tantangan tersendiri karena sifat berita yang tidak terstruktur. Penelitian ini menggunakan model word2vec untuk merepresentasikan berita ke dalam bentuk vektor, dan metode Long Short-Term Memory untuk klasifikasinya. Kami memanfaatkan 11.854 dataset berita saham dengan 750 lebih kode saham yang berasal dari Indonesia pada rentang tahun 2020-2021. Berdasarkan 6 model eksperimen yang dilakukan: (1) penggunaan data; (2) baseline; (3) arsitektur; (4) penambahan data; (5) hyperparameter; dan (6) penggunaan data 100 saham tertinggi, temuan memperlihatkan bahwa berita yang signifikan mempengaruhi pergerakan harga saham adalah berita yang terbit pada 6-7 hari sebelumnya. Hasil F1-Score terbaik yang didapatkan yaitu 0,32 untuk kelas harga naik dan 0,43 untuk kelas harga turun menggunakan data 6 hari setelah berita terbit, arsitektur LSTM 1-layer sebagai layer input, batch size 64, dan optimizer adam.

Kata Kunci: Stock Prediction, LSTM, Word2Vec

PREDIKSI PERGERAKAN HARGA SAHAM BERDASARKAN BERITA MENGUNAKAN METODE *LONG SHORT-TERM MEMORY*

Arranged by

Fenny Feronika Maharani — fenny@upi.edu

1703015

ABSTRACT

The stock market has a complex and dynamic behavior, so making investment decisions in the stock market is a difficult task. Public information in the form of news can be used as data to predict stock price movements, but prediction of stock price movements based on news has its own challenges because of the unstructured nature of news. This study uses the word2vec model to represent news in vector form, and the Long Short-Term Memory method for classification. We utilize 11,854 stock news datasets with more than 750 stock codes originating from Indonesia in the range of 2020-2021. Based on 6 experimental models: (1) data usage; (2) baselines; (3) architecture; (4) addition of data; (5) hyperparameters; and (6) the use of data from the highest 100 stocks, the findings show that news that significantly affects stock price movements is news published on the previous 6-7 days. The best F1-Score results obtained are 0.32 for the rising price class and 0.43 for the descending price class using data 6 days after the news was published, 1-layer LSTM architecture as the input layer, batch size 64, and the adam optimizer. Keyword: *Stock Prediction*, LSTM, Word2Vec

KATA PENGANTAR

Puji dan syukur selalu tercurahkan atas kehadiran Allah SWT karena hanya kebesaran-Nya dan kehendak-Nya serta karunia-Nya penyusunan skripsi yang berjudul “Prediksi Pergerakan Harga Saham Berdasarkan Berita Menggunakan Metode *Long Short-Term Memory*” ini dapat selesai.

Penyusunan skripsi ini ditujukan untuk memenuhi dan melengkapi salah satu syarat untuk mendapatkan gelar sarjana komputer atas jenjang S1 pada Program Studi Ilmu Komputer Universitas Pendidikan Indonesia.

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 agar tidak terjadi kesalahan yang sama di kemudian hari dan dapat meningkatkan kualitas ke tahap yang lebih baik.

Bandung, Maret 2022

Penulis

UCAPAN TERIMAKASIH

Alhamdulillah, segala puji bagi Allah SWT yang telah mengkaruniakan berkah dan kasih sayang-Nya sehingga atas izin-Nya penulis dapat menyelesaikan skripsi ini yang berjudul “Prediksi Pergerakan Harga Saham Berdasarkan Berita Menggunakan Metode *Long Short-Term Memory*”.

Penulis menyusun skripsi ini dalam rangka memenuhi salah satu persyaratan untuk mencapai gelar sarjana (S1) pada Program Studi Ilmu Komputer, FPMIPA Universitas Pendidikan Indonesia.

Penulis tentunya tidak dapat menyelesaikan penelitian ini tanpa bantuan dan dorongan dari pihak-pihak yang telah membantu baik secara langsung ataupun tidak langsung. Maka dari itu pada kesempatan ini pula, penulis mengucapkan terima kasih kepada:

1. Allah SWT. yang telah memberikan rahmat dan karunia-Nya sehingga penulis dapat melewati proses dan menyelesaikan skripsi dengan baik.
2. Kedua orang tua penulis, Bapak Rudi Sunarko dan Ibu Siti Maemunah juga kakak kandung penulis Ardi Hardiyanto yang tanpa henti-hentinya memberikan doa dan dukungan, baik itu dukungan moral, material maupun spiritual sehingga dapat memotivasi penulis dalam menyelesaikan skripsi ini.
3. Bapak Dr. Yudi Wibisono, M.T. selaku pembimbing I yang telah bersedia memberikan masukan, saran, dan kritik yang membangun serta meluangkan waktu, tenaga, dan pikiran dalam memberikan pengarahan dalam penelitian maupun penulisan skripsi ini.
4. Bapak Herbert Siregar, M.T. selaku pembimbing II yang telah memberikan saran dan arahan kepada penulis selama penulis menjalankan proses penyelesaian penelitian dan penulisan skripsi.
5. Ibu Rani Megasari M.T., selaku Ketua Program Studi Ilmu Komputer FPMIPA Universitas Pendidikan Indonesia.
6. Seluruh dosen Ilmu Komputer yang telah membimbing dan memberikan ilmu yang bermanfaat pada penulis selama masa kuliah dan staf administrasi Program Studi Ilmu Komputer yang telah memberikan informasi akademik selama masa perkuliahan

7. Sahabat di kelas C 2017 yang telah bersama-sama menikmati indahny masa perkuliahan.
8. Tidak lupa kepada Nur ‘Aisyah Nadiyah, Rantty Gantini, Rizal Purnomo, Achmad Firditama dan Ade Prianto terimakasih telah menjadi teman berdiskusi dan berbagi keluh kesah serta motivasi selama proses penyusunan skripsi.
9. Kepada Muhammad Roychan Meliaz yang senantiasa memberikan dukungan moral dan membantu penulis dalam proses penelitian dan penulisan skripsi hingga selesai.
10. Kepada diri sendiri yang masih bisa semangat dan dapat menyelesaikan skripsi ini dengan baik.
11. Semua pihak yang tidak bisa disebutkan satu persatu yang telah memberi arti dan dukungan pada penulis

Tidak ada kata-kata yang dapat menggambarkan rasa terima kasih penulis atas semua dukungan yang telah diberikan semoga atas kebaikan dan keikhlasan kepada penulis dalam membantu menyelesaikan skripsi dapat di balas oleh Allah SWT dan senantiasa diberikan kesehatan oleh Allah SWT.

Bandung, Maret 2022

Fenny Feronika Maharani
1703015

DAFTAR ISI

PERNYATAAN.....	iii
ABSTRAK.....	iv
ABSTRACT.....	v
KATA PENGANTAR	vi
UCAPAN TERIMAKASIH.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Penelitian	4
1.6 Sistematika Penulisan	5
BAB II KAJIAN PUSTAKA	7
2.1 Penelitian terkait.....	7
2.2 Pengertian Saham.....	8
2.3 Jenis-jenis Saham	9
2.4 Return Saham	10
2.5 Resiko Saham.....	11
2.6 Pasar Modal.....	12
2.7 Word Embedding	13
2.8 Word2Vec	13
	ix

2.9 Baseline	14
2.10 Recurrent Neural Network	14
2.11 Long Short-Term Memory	15
2.12 Hyperparameter	21
2.13 Evaluasi Model.....	21
BAB III METODOLOGI PENELITIAN	23
3.1 Desain Penelitian.....	23
3.2 Alat dan Bahan Penelitian.....	26
3.2.1 Alat Penelitian	26
3.2.2 Bahan Penelitian.....	27
BAB IV HASIL DAN PEMBAHASAN	28
4.1 Hasil	28
4.1.1 Pembuatan Dataset	28
4.1.2 Perancangan Model Komputasi.....	36
4.1.3 Skenario Eksperimen.....	40
4.1.4 Eksperimen.....	47
4.2 Pembahasan.....	53
BAB V KESIMPULAN	57
5.1 Kesimpulan	57
5.2 Saran.....	58
DAFTAR PUSTAKA	59

DAFTAR GAMBAR

Gambar 2. 1 Arsitektur RNN	15
Gambar 2.2 Arsitektur LSTM.....	17
Gambar 2. 3 Komponen LSTM	18
Gambar 2. 4 Operasi Forget gate	18
Gambar 2. 5 Operasi Input gate	19
Gambar 2. 6 Operasi <i>Cell state</i>	20
Gambar 2. 7 Operasi <i>Output gate</i>	20
Gambar 3.1 Desain Penelitian.....	23
Gambar 4.1 Diagram Alur Proses Scraping Data Berita Saham.....	29
Gambar 4.2 Laman <i>Website</i> ww.iqplus.info	30
Gambar 4.3 Inspect Element Laman <i>Website</i> www.iqplus.info	30
Gambar 4.4 Diagram Alur Pelabelan Data	34
Gambar 4.5 Model Komputasi Sistem.....	36
Gambar 4. 6 Visualisasi Proses Vektorisasi.....	39

DAFTAR TABEL

Tabel 2. 1 <i>Confusion Matrix</i>	22
Tabel 4. 1 Hasil <i>Scraping</i> Data Berita Saham	31
Tabel 4. 2 Contoh Data Harga Saham.....	33
Tabel 4.3 Contoh Sebagian Dataset	35
Tabel 4. 4 Hasil Praproses Data	37
Tabel 4. 5 Hasil <i>Tokenization</i>	38
Tabel 4. 6 Hasil Vektorisasi	39
Tabel 4.7 Visualisasi Dataset	41
Tabel 4. 8 Arsitektur yang digunakan	44
Tabel 4. 9 Daftar 100 Index Saham Dengan Pergerakan Tertinggi	45
Tabel 4. 10 Hasil Eksperimen Baseline	47
Tabel 4.11 Hasil Eksperimen Arsitektur	48
Tabel 4. 12 Hasil Eksperimen Penambahan Data	50
Tabel 4. 13 Hasil Eksperimen Batch Size	50
Tabel 4. 14 Hasil Eksperimen Optimizer	51
Tabel 4.15 Hasil Eksperimen Besar Padding	52
Tabel 4. 16 Hasil Eksperimen Data 100 Pergerakan Saham Tertinggi.....	53
Tabel 4. 17 Contoh Hasil Prediksi	53
Tabel 4. 18 Hasil Terbaik Dari Eksperimen Pelatihan Baseline	54
Tabel 4. 19 Hasil Terbaik dari Eksperimen Arsitektur	55
Tabel 4. 20 Hasil Terbaik Dari Eksperimen Hyperparameter Batch Size	55
Tabel 4.21 Hasil Terbaik Dari Eksperimen Hyperparameter Optimizer	56

DAFTAR PUSTAKA

- Admodjo, B. S. (2021). *Pre-Trained Word2vec Indonesia*. Kaggle. <https://www.kaggle.com/datasets/bhimantoros/pretrained-word2vec-indonesia>
- Aggarwal, C. C. (2014). Educational and software resources for data classification. *Data Classification: Algorithms and Applications*, 657–665. <https://doi.org/10.1201/b17320>
- Alostad, H., & Davulcu, H. (2016). Directional prediction of stock prices using breaking news on twitter. *Proceedings - 2015 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, WI-IAT 2015*, 1, 523–530. <https://doi.org/10.1109/WI-IAT.2015.82>
- Azlina, N. (2009). Pengaruh the Monday Effect Terhadap Return Saham Jii Di Bursa Efek Indonesia. *Pekbis Jurnal*, 1(1), 26–35.
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *ArXiv Preprint ArXiv:1406.1078*.
- Claesen, M., & De Moor, B. (2015). *Hyperparameter Search in Machine Learning*. 10–14. <http://arxiv.org/abs/1502.02127>
- Di Persio, L., & Honchar, O. (2016). Artificial neural networks approach to the forecast of stock market price movements. *International Journal of Economics and Management Systems*, 1.
- Eduardus Tendelilin. (2001). Analisis Investasi dan Manajemen Portofolio, Edisi I, cet. I. Yogyakarta, BPFE.
- Emmanuel Ameisen. (2018). *Always start with a stupid model, no exceptions*. Medium. <https://blog.insightdatascience.com/always-start-with-a-stupid-model-no-exceptions-3a22314b9aaa>
- Ghannay, S., Benoit Favre, Yannick Esteve, & Nathalie Camelin. (2015). *Word Embeddings Evaluation and Combination*. 1, 1–14.

- Graves, A. (2012). Connectionist temporal classification. In *Supervised Sequence Labelling with Recurrent Neural Networks* (pp. 61–93). Springer.
- Gunawan, W. (2021). *Dataset Saham IDX*. GitHub. <https://github.com/wildangunawan/Dataset-Saham-IDX>
- Hochreiter, S., & Schmidhuber, J. (1997a). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
- Hochreiter, S., & Schmidhuber, J. (1997b). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Husnan, S. (2001). Dasar-Dasar Teori Portofolio dan Analisis Sekuritas Edisi Ketiga. Yogyakarta: UPP AMP YKPN.
- Iqplus. (2020). *Stock News*. Iqplus.Info. http://www.iqplus.info/news/stock_news/go-to-page,0.html
- Johnson, F., & Kumar Gupta, S. (2012). Web Content Mining Techniques: A Survey. *International Journal of Computer Applications*, 47(11), 44–50. <https://doi.org/10.5120/7236-0266>
- Karpathy, A. (2015). *Multi-layer recurrent neural networks (lstm, gru, rnn) for character-level language models in torch*.
- Levy, O., & Goldberg, Y. (2014). Neural word embedding as implicit matrix factorization. *Advances in Neural Information Processing Systems*, 3(January), 2177–2185.
- Li, D., Hasanaj, E., & Li, S. (2020). 3 – Baselines. Machine Learning Departement Carnegie Melon University. <https://blog.ml.cmu.edu/2020/08/31/3-baselines/>
- Li, X., Li, Y., Yang, H., Yang, L., & Liu, X.-Y. (2019). DP-LSTM: Differential Privacy-inspired LSTM for Stock Prediction Using Financial News. *NeurIPS*, 1–9. <http://arxiv.org/abs/1912.10806>
- Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). A critical review of recurrent neural networks for sequence learning. *ArXiv Preprint ArXiv:1506.00019*.

- Ma, L., & Zhang, Y. (2015). Using Word2Vec to process big text data. *Proceedings - 2015 IEEE International Conference on Big Data, IEEE Big Data 2015*, 2895–2897. <https://doi.org/10.1109/BigData.2015.7364114>
- Marketstack. (2020). *Real-Time, Intraday & Historical Market Data API*. Marketstack. www.marketstack.com
- Mikolov, T., Corrado, G., Chen, K., & Dean, J. (n.d.). *Vector Space*. 1–12.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems* (Vol. 26). Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf>
- Nurdin, A., Anggo, B., Aji, S., Bustamin, A., & Abidin, Z. (2020). PERBANDINGAN KINERJA WORD EMBEDDING WORD2VEC, GLOVE, DAN FASTTEXT PADA KLASIFIKASI TEKS. *Jurnal TEKNOKOMPAK*, 14(2), 74.
- Prabowo, & Tommy. (2000). Dissemination of information di Pasar Modal. *Media Akuntansi*, 10.
- Rahardja, U., Lutfiani, N., & Rahmawati, R. (2018). Persepsi Mahasiswa Terhadap Berita Pada Website APTISI. *Jurnal Ilmiah SISFOTENIKA*, 8. <https://doi.org/10.30700/jst.v8i2.400>
- Rahmadzani, R. F. (2021). *Cara Kerja Long Short-Term Memory (LSTM) | Catatan Penelitian #11*. <https://rifqifai.com/cara-kerja-long-short-term-memory-lstm/>
- Ramadhan, N. G., & Atastina, I. (2021). *Neural Network on Stock Prediction using the Stock Prices Feature and Indonesian Financial News Titles*. 54–63. <https://doi.org/10.34818/ijoiect.v7i1.544>
- Rather, A. M., Agarwal, A., & Sastry, V. N. (2015). Recurrent neural network and

- a hybrid model for prediction of stock returns. *Expert Systems with Applications*, 42(6), 3234–3241.
- Rechtschaffen, A. N. (2019). *Capital Markets, Derivatives, and the Law* (3rd ed.). Oxford University Press.
- Saad, E. W., Prokhorov, D. V., & Wunsch, D. C. (1998). Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks. *IEEE Transactions on Neural Networks*, 9(6), 1456–1470.
- Saputro, I. W., & Sari, B. W. (2020). Uji Performa Algoritma Naïve Bayes untuk Prediksi Masa Studi Mahasiswa. *Creative Information Technology Journal*, 6(1), 1. <https://doi.org/10.24076/citec.2019v6i1.178>
- Sawidji, W. (1996). Cara sehat investasi di pasar modal. *PT. Jurnalindi Aksara Grafika, Jakarta*.
- Septyanto Dihin. (2013). Faktor-Faktor Yang Mempengaruhi Investor Individu Dalam Pengambilan Keputusan Investasi Sekuritas Di Bursa Efek Indonesia (BEI). *Jurnal Ekonomi*, 4(2), 90–101.
- Silvani, Y. (2006). Pengaruh Faktor-Faktor Fundamental Keuangan Terhadap Harga Saham Pada Perusahaan Listing di BEJ. *Skripsi Universitas Riau, Pekanbaru*.
- Sjahrir. (1995). *Tinjauan Pasar Modal*. PT Gramedia Pustaka Utama.
- Sulistio, & Helen. (2005). Pengaruh Informasi Akuntansi dan Non Akuntansi Terhadap Initial Return: Studi pada Perusahaan yang Melakukan Initial Public Offering di Bursa Efek Jakarta. *Simposium Nasional Akuntansi VIII*, 87--99.
- Sunariyah. (2006). *Pengantar Pengetahuan PasarModal Edisi ke 5* (5th ed.). ANDI.
- Syahrul, M. A. Z. (2000). *Kamus Lengkap Ekonomi: Istilah-istilah Akuntansi, Keuangan, dan Investasi*. Bandung: Citra Harta Prima.
- Tandelilin, E. (2001). *Analisis investasi dan manajemen portofolio*. Yogyakarta: Bpfe.

- Trisnawati, W. (2013). Pengaruh arus kas operasi, investasi dan pendanaan serta laba bersih terhadap retur saham. *Jurnal Ilmu Dan Riset Akuntansi*, 1(1), 77–92.
- Wahyudi, S. (2003). Pengukuran Return Saham. *Jurnal Ekonomi. Suara Merdeka*.
- Walter, W., H, Kin-Yip, Liu, Raymond, W.-M., & Tracy, W. K. (2013). The relation between news and stock price jump: an analysis based on neural network. *Modelling and Simulation Society of Australia and New Zealand Inc.*
- Zuhrotun, & Baridwan. (2005). Pengaruh Pengumuman Peringkat Terhadap Kinerja Obligasi. *Prosiding. Simposium Nasional Akuntansi VIII. Solo*.