

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pasar saham memiliki peranan penting dalam roda perekonomian dan merupakan alternatif bagi pendanaan dan sarana berinvestasi. Pasar saham memiliki perilaku yang kompleks dan sifat yang tidak stabil oleh karena itu pengambilan keputusan di pasar saham adalah tugas yang sangat sulit dan penting. Dalam menentukan waktu yang tepat untuk membeli atau menjual saham para investor harus menemukan cara agar bisa mencapai keuntungan terbaik dari investasi mereka. Kebutuhan yang paling mendasar bagi seorang investor adalah informasi, agar hasil yang diperoleh sesuai dengan yang diharapkan maka informasi yang dimiliki oleh investor harus lengkap, akurat dan tepat waktu sehingga investor dapat melakukan pengambilan keputusan secara rasional. Brun (dalam Septyanto Dihin, 2013) menjelaskan bahwa bagi investor, informasi merupakan signals yang berfungsi sebagai stimulus. Informasi yang beredar di publik seringkali dimanfaatkan oleh investor karena prospek suatu perusahaan di masa depan terdapat dalam informasi tersebut. (Sulistio & Helen, 2005; Zuhrotun & Baridwan, 2005).

Salah satu informasi yang beredar di publik adalah berita. Berita merupakan kumpulan informasi yang dapat berupa teks, ucapan maupun gambar yang dibutuhkan sebagai penambah dan pembaruan wawasan atau informasi yang dimiliki oleh manusia (Rahardja et al., 2018). Berita adalah salah satu informasi yang dapat digunakan dalam melakukan keputusan investasi, maka dari itu berita memiliki pengaruh terhadap perilaku pasar saham. Secara tidak langsung berita akan mempengaruhi harga saham karena berita mempengaruhi keputusan investor dan keputusan investor akan mempengaruhi harga saham. Berita memiliki sifat tidak terstruktur, dalam pengolahannya diperlukan proses ekstraksi pola berupa informasi dan pengetahuan yang berguna dari sumber data berita itu sendiri.

Dalam menanggapi informasi atau berita yang beredar investor memiliki respon tersendiri namun memiliki kemampuan analisis yang terbatas dalam

menginterpretasikan informasi yang didapat (Septyanto Dihin, 2013). Hal itu menyebabkan investor bertindak secara *naïve, irrational, dan unsophisticated* (tidak canggih) (Prabowo & Tommy, 2000). Seorang investor cenderung menentukan keputusannya berdasarkan pada rumor, isu, spekulatif, dan berperilaku *mass behaviour, impulsivity, loss-control, dan impatience* (Sjahrir, 1995). Untuk membantu investor dalam melakukan keputusan investasi, prediksi pergerakan harga saham dapat dijadikan solusi karena dapat memberi investor wawasan yang kuat tentang perilaku pasar saham untuk menghindari resiko investasi.

Prediksi pergerakan harga saham dapat dilakukan berdasarkan berita dan bidang tersebut merupakan bidang penelitian yang menarik dan memiliki banyak sekali tantangan karena sifat berita yang tidak terstruktur. Penambangan berita dapat didefinisikan sebagai proses penggalian pola tersembunyi, berguna dan berpotensi untuk memperoleh pengetahuan yang tidak diketahui dari data berita. Untuk mengatasi sifat berita yang tidak terstruktur maka perlu dilakukan penambangan teks, penambangan teks juga dikenal sebagai dalam Knowledge Discovery in Text (KDT). Dalam penelitiannya Wang Walter menyelidiki hubungan antara berita keuangan dan perilaku pasar saham menggunakan kausalitas granger dan mengungkapkan bahwa ada hubungan antara sentimen berita dan perilaku pasar saham (Walter et al., 2013).

Dalam melakukan prediksi saham telah digunakan berbagai algoritma dari pembelajaran mesin seperti *neural networks, genetic algorithms, support vector machine*, dan lainnya. *Recurrent neural networks* (RNN) adalah model yang andal untuk memproses data sekuensial seperti suara, data deret waktu atau bahasa (Lipton et al., 2015). Beberapa desain RNN digunakan untuk memprediksi, diantaranya adalah Long Short-Term Memory (LSTM) (Rather et al., 2015; Saad et al., 1998). LSTM menurut Hochreiter & Schmidhuber adalah salah satu arsitektur RNN yang paling sukses. Arsitektur ini telah digunakan secara luas dan mampu mencapai beberapa hasil terbaik jika dibandingkan dengan metode lain (Graves, 2012). (Di Persio & Honchar, 2016) membandingkan kinerja LSTM dan MLP dengan metode yang mereka usulkan sendiri yaitu kombinasi wavelet dan

Convolutional Neural Networks, hasilnya metode tersebut mengungguli LSTM dan MLP tetapi memiliki hasil yang sangat mirip dengan LSTM.

Untuk mendapatkan hasil yang optimal tidak cukup hanya menggunakan algoritma pembelajaran mesin saja, namun dalam bidang pemrosesan teks diperlukan teknik representasi kata ke dalam vektor. *Feature Engineering* dalam data tekstual menjadi sangat penting karena data tekstual memiliki karakteristik yang tidak terstruktur (Nurdin et al., 2020). *Word embedding* adalah salah satu strategi *feature engineering* yang memetakan setiap kata dalam dokumen ke dense vector, di mana sebuah vector merepresentasikan proyeksi kata di dalam ruang vector. Sehingga dalam teknik *word embedding* ini dapat menangkap makna semantic dan sintatik kata, karena posisi kata tersebut dipelajari dari teks berdasarkan kata-kata di sekitarnya (Nurdin et al., 2020). Pada tahun 2013 Mikolov et al., n.d. menemukan salah satu algoritma *word embedding* yaitu Word2Vec. Word2Vec memetakan setiap kata di dalam dokumen ke dalam vektor, mengandalkan informasi lokal dari bahasa. Semantik yang dipelajari dari kata tertentu dipengaruhi oleh kata-kata sekitarnya. Model ini mendemonstrasikan kemampuan untuk mempelajari pola linguistik sebagai hubungan linear antar vektor kata (Nurdin et al., 2020).

Berdasarkan uraian diatas, maka penulis ingin membuat penelitian untuk merancang model prediksi harga saham menggunakan berita dengan mengimplementasikan metode *word embedding* dengan algoritma Word2Vec dan kemudian digabungkan dengan metode *machine learning Long Short-Term Memory* (LSTM). Metode ini dipilih berdasarkan studi literatur dari penelitian terkait yang menunjukkan bahwa word embedding dengan algoritma Word2Vec mengungguli algoritma tradisional dalam ekstraksi fitur teks. Metode *machine learning* LSTM juga dibuktikan memiliki hasil yang sama bagusnya dengan metode yang diusulkan pada penelitian terkait perbandingan kinerja antara LSTM dan MLP yang telah dikombinasikan dengan metode lain. Harapannya dari penelitian yang dilakukan oleh penulis dapat dihasilkan model yang baik dan dapat membantu investor agar dapat memaksimalkan keuntungan, meminimalkan biaya dan resiko

seiring meningkatnya kesadaran investor terhadap pergerakan harga saham yang mengarah kepada keputusan investasi yang kuat.

1.2 Rumusan Masalah

Berdasarkan permasalahan yang terdapat pada latar belakang masalah, maka rumusan permasalahan sebagai berikut:

1. Bagaimana desain rancangan model untuk prediksi harga saham berdasarkan berita?
2. Bagaimana hasil kinerja model prediksi harga saham yang telah dirancang?

1.3 Tujuan Penelitian

Mengacu pada rumusan masalah yang telah diperoleh sebelumnya, maka tujuan dilakukannya penelitian ini adalah:

1. Merancang model prediksi harga saham berdasarkan berita dengan menggunakan metode *Long Short-Term Memory* dan metode *word embedding* dengan algoritma Word2Vec.
2. Mengevaluasi hasil kinerja dan model prediksi harga saham berdasarkan berita yang telah dibangun berdasarkan tingkat F1-Score.

1.4 Manfaat Penelitian

Manfaat penelitian ini untuk merancang dan menganalisis model yang dapat digunakan dalam memprediksi pergerakan harga saham berdasarkan berita. Penelitian ini juga bermanfaat untuk memahami pengaruh berita saham terhadap pergerakan harga saham. Pada penelitian ini juga dilakukan evaluasi rancangan model yang dibuat agar mendapatkan model yang terbaik.

1.5 Batasan Penelitian

Agar pembahasan penelitian ini lebih terarah maka diberikan batasan-batasan masalah sebagai berikut:

1. Berita yang digunakan adalah berita saham yang diperoleh dari portal berita iqplus.info pada kurun waktu tahun 2018 – 2021.

2. Data numerik harga saham diperoleh dari marketstack.com dan dataset harga saham milik Wildan Gunawan pada kurun waktu 2018-2021.
3. Model dibuat untuk memprediksi harga saham dalam kurun waktu 3,4,5,6, dan 7 hari ke depan.

1.6 Sistematika Penulisan

Sistematika penulisan pada penelitian ini terdiri dari:

BAB I PENDAHULUAN

Bab pendahuluan ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, metode yang digunakan serta sistematika penulisan.

BAB II KAJIAN PUSTAKA

Bab ini berisi landasan teori serta informasi yang dijadikan sumber yang berkaitan dengan saham, *word embedding* Word2Vec, dan algoritma *Long Short-Term Memory* yang di terapkan dalam penelitian.

BAB III METODOLOGI PENELITIAN

Bab ini akan memberikan desain penelitian atau alur penelitian dari mulai awal penelitian identifikasi masalah dari bab sebelumnya sampai berhasil menciptakan sebuah model pada akhir penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini membahas proses pengumpulan data untuk proses penelitian dan hasil eksperimen yang telah dilakukan selama penelitian. Semua pertanyaan mengenai masalah yang diangkat dalam tema skripsi dibahas pada bab ini, diantaranya adalah tentang pengembangan model, implementasi sistem, desain eksperimen dan analisa.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari penelitian tentang prediksi harga saham berdasarkan berita yang telah dilakukan penulis. Penulis juga memberi saran bagi peneliti selanjutnya dari hasil penelitian yang telah dilakukan.