



### **Hak cipta dan penggunaan kembali:**

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

### **Copyright and reuse:**

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

## **BAB II**

### **LANDASAN TEORI**

#### **2.1. Harga Saham**

Menurut Anoraga, harga saham adalah uang yang dikeluarkan untuk memperoleh bukti penyertaan atau kepemilikan suatu perusahaan. Calon investor dapat mengetahui seberapa jauh manajemen telah berhasil mengelola perusahaan atas nama pemegang saham dengan melihat harga saham, dengan kata lain harga saham merupakan ukuran indeks prestasi perusahaan (Gitaya, 2016). Harga saham terbentuk melalui mekanisme permintaan dan penawaran di pasar modal. Apabila suatu saham mengalami kelebihan permintaan, maka harga saham cenderung naik. Sebaliknya, apabila kelebihan penawaran maka harga saham cenderung turun (Rachmawati, 2017).

Saat ini, sejumlah besar informasi berharga yang berkaitan dengan pasar keuangan tersedia di berbagai media seperti situs *web*, Twitter, Facebook, *blog*, dan lainnya. Secara umum harga saham tergantung pada dua faktor. Salah satunya adalah faktor fundamental (*fundamental factor*) dan faktor lainnya adalah faktor teknis (*technical factor*). Faktor fundamental lebih mengutamakan ketergantungan pada data statistik perusahaan, seperti mencakup laporan, status keuangan perusahaan, neraca, dividen, dan kebijakan perusahaan yang sahamnya harus diperhatikan. Faktor teknis mencakup parameter kuantitatif seperti indikator tren, naik turun harian, nilai tertinggi dan terendah dalam sehari, volume saham, indeks,

rasio *put / call*, dll. Dalam faktor teknis, harga historis dipertimbangkan untuk melakukan prediksi. Harga historis perusahaan mudah dicari dengan mengunduh datanya dari situs *web*. Berbagai metode indikator *stock level* tersedia untuk menghitung nilai saham, diantaranya adalah *Moving Average*, *Stochastic RSI (Relative-Strength Index)*, *Bollinger bands*, *Accumulation – Distribution*, dan *Typical Point (pivot point)* (Bharathi & Geetha, 2017).

Sedangkan menurut Arifin, salah satu faktor yang mempengaruhi harga saham adalah *news* dan rumor. Semua berita yang beredar di masyarakat yang menyangkut beberapa hal baik itu masalah ekonomi, sosial, politik keamanan, hingga berita seputar *reshuffle* kabinet (Wahyuni, 2015). Para investor bisa memprediksi seberapa kondusif keamanan negeri ini dengan adanya berita-berita tersebut sehingga kegiatan investasi dapat dilaksanakan dan berdampak pada pergerakan harga saham di bursa.

## **2.2. Moving Average**

*Moving average* mengacu pada praktik pembelian dan penjualan secara sistematis setiap kali harga melewati rata-ratanya. Idennya adalah harga bergerak dalam tren sehingga pada setiap titik waktu, harga berada dalam *uptrend* (tren naik) atau *downtrend* (tren turun). *Uptrend* didefinisikan sebagai periode kenaikan harga dan *downtrend* didefinisikan sebagai periode penurunan harga. Ketika harga melewati atau memotong nilai rata-rata dari bawah, karena harga saat ini lebih tinggi dari harga yang lebih tua, harga dikatakan dalam tren *uptrend* dan *buy signal*

(sinyal beli) terjadi. Demikian pula, ketika harga melewati atau memotong rata-rata dari atas, karena harga saat ini lebih rendah dari harga yang lebih tua, harga dikatakan dalam tren *downtrend* dan *sell signal* (sinyal jual) terjadi. Respons yang mengikuti *buy signal* adalah dengan *action buy* (membeli) dan respons yang mengikuti *sell signal* adalah dengan *action sell* (menjual). Jika perubahan tingkat harga di antara *buy signal* dan *sell signal* cukup untuk menutup biaya, *moving average* berarti menguntungkan. Sebaliknya, jika perubahan tingkat harga di antara *buy signal* dan *sell signal* tidak cukup untuk menutupi biaya, *moving average* berarti memberikan kerugian (Toms, 2011).

Keuntungan utama dari indikator level saham *moving average* adalah ditawarkannya garis yang halus dan juga membantu mengurangi jumlah *noise* pada grafik harga dibandingkan dengan tingkat indikator lainnya. Untuk saham, periode waktu yang umum untuk *moving average* adalah 5 hari, 10 hari, 15 hari, 21 hari, 50 hari, 100 hari dan 200 hari. *Moving average* dihitung dengan menambahkan harga *closing* kemudian membagi total harga *closing* dengan jumlah periode waktu (Bharathi & Geetha, 2017). Berikut ini penjelasan rumus *moving average*.

$$F_t = \frac{A_{t-1} + A_{t-2} + A_{t-3} + \dots + A_{t-n}}{n}$$

**Rumus 2. 1. Moving Average**

$F_t$  = Prediksi periode mendatang

$A_{t-1}$  = Kejadian aktual di masa lalu hingga periode 'n'

$n$  = Jumlah periode yang akan dirata-rata

### 2.3. *Text Mining*

*Text mining* secara umum adalah teori tentang pengolahan koleksi dokumen dalam jumlah besar yang ada dari waktu ke waktu dengan menggunakan beberapa analisis. Tujuan pengolahan teks tersebut adalah mengetahui dan mengekstrak informasi yang berguna dari sumber data dengan identifikasi dan eksplorasi pola. Pada kasus *text mining*, sumber data yang digunakan adalah kumpulan atau koleksi dokumen tidak terstruktur dan memerlukan adanya pengelompokan untuk diketahui informasi sejenis (Somantrii, Wiyono, & Dairoh, 2016). Untuk memberikan solusi, *text mining* mengadopsi dan mengembangkan banyak teknik dari bidang lain, seperti *data mining*, *information retrieval*, statistik dan matematika, *machine learning*, *linguistic*, *natural language processing*, dan *visualization*. Proses *text mining* ini dimulai dari pengumpulan data kemudian dilakukan proses *pre-processing* yang didalamnya terdapat proses *categorization*, *feature/term extraction*, setelah melalui proses tersebut dokumen masuk ke dalam proses pemberian *keyword*, *labeled* dan *time-stamped*, dan terakhir adalah proses *core mining operations* dan *presentation* (Nurhadi, 2015).

### 2.4. Analisis Sentimen

Analisis sentimen yang juga disebut dengan *opinion mining* merupakan bidang studi yang menganalisis pendapat, penilaian, evaluasi, sentimen, sikap dan emosi terhadap entitas seperti produk, jasa, organisasi, individu, peristiwa, topik dan atribut lainnya. *Big data* dari situs media sosial seperti Twitter dapat digunakan untuk analisis. (Chakra, Gupta, & Kumar, 2018).

Dua metode utama analisis sentimen adalah metode *lexicon-based* dan pendekatan *machine learning* dimana keduanya mengandalkan *bag-of-words*. Pada *machine learning* metode *supervised*, pengklasifikasi menggunakan *unigrams* sebagai fitur. Dalam metode *lexicon-based*, *unigram* yang ditemukan dalam *lexicon* diberi skor *polarity*, skor *polarity* keseluruhan teks kemudian dihitung sebagai jumlah dari polaritas *unigram* (Kolchyna, Souza, Treleaven, & Aste, 2015).

## 2.5. Twitter

Twitter adalah layanan jejaring sosial dan situs *microblogging* dimana *user* bisa mengirim sesuatu dan saling berinteraksi dengan pesan teks yang dinamakan '*tweet*'. Setiap *tweet* dibatasi dengan 280 karakter. Menurut Statista, Twitter bukanlah media sosial yang paling populer jika dilihat dari segi banyaknya pengguna yang aktif. Peringkat pertama diduduki oleh Facebook sedangkan Twitter hanya menempati posisi kedua belas (*Most popular social networks worldwide as of April 2018, ranked by number of active users (in millions)*, 2018).

Analisis sentimen Twitter telah digunakan untuk mempelajari *behavioral finance*, khususnya, efek sentimen yang didorong dari media sosial pada keputusan keuangan dan ekonomi. Tetapi *tweet* yang diambil hanya dari akun Twitter "Verified" atau "Terverifikasi". Akun *verified* di Twitter menunjukkan bahwa akun itu adalah kepentingan umum dan akun itu asli (Tabari, Praneeth, & Biswas, 2018). Penelitian kali ini akun *verified* dibatasi menjadi akun *outlet news*. Mengambil informasi berita akun *outlet news* dari *platform* Twitter saja sudah cukup karena (a)

Twitter telah menjadi saluran populer untuk pelaporan berita oleh media *mainstream* (b) sejumlah besar *tweet* memungkinkan kita untuk mendapatkan wawasan tentang aspek temporal dari liputan berita dengan lebih baik dibandingkan dengan artikel berita (c) sifat *tweet* yang tidak disensor yang dikumpulkan dengan Twitter Streaming API (Ali, Stratmann, & Park, 2018). Akun *official news* yang menjadi sumber data adalah *outlet news* yang biasa menjadi sumber berita keuangan, yaitu Wall Street Journal, Market Watch, Forbes, Reuters, dan CNBC (*Consumer News and Business Channel*) (Kirange & Deshmukh, 2016).

## 2.6. Facebook, Inc

Facebook adalah perusahaan besar yang terkenal dengan layanan jejaring sosialnya. Didirikan pada tahun 2004, misi Facebook adalah memberikan masyarakat kekuatan untuk membuat komunitas dan menyatukan dunia. Facebook juga memiliki anak perusahaan yang tidak kalah besarnya dengan Facebook sendiri, seperti Instagram dan Whatsapp. Pada bulan April tahun 2018, CNN Money melaporkan saham Facebook turun sampai 10% setelah beredarnya skandal data Cambridge Analytica (*Facebook is the worst of the Big Tech stocks*, 2018).

## 2.7. R

R adalah *software environment* gratis untuk komputasi dan grafis statistik. R terdiri dari bahasa *run-time environment*, grafik, *debugger*, akses ke fungsi sistem tertentu, dan kemampuan untuk menjalankan program yang disimpan dalam file

*script*. Inti dari R adalah bahasa komputer yang ditafsirkan yang memungkinkan percabangan dan pengulangan serta pemrograman modular menggunakan fungsi. Sebagian besar fungsi yang dapat dilihat pengguna dalam R ditulis dalam R. Memungkinkan pengguna untuk berinteraksi dengan prosedur yang ditulis dalam bahasa C, C ++, atau FORTRAN untuk efisiensi. Distribusi R berisi fungsionalitas untuk sejumlah besar prosedur statistik. Di antaranya adalah linear dan *generalized linear models*, *nonlinear regression models*, *time series analysis*, *classical parametric* dan *nonparametric tests*, *clustering* dan *smoothing* (Hornik, 2017).

## 2.8. Python

Python adalah bahasa pemrograman tingkat tinggi yang dibuat oleh Guido van Rossum. Bahasa ini sangat populer dengan kemudahannya untuk membaca *code* dan garis *code* yang ringkas. Menggunakan penggenangan *white spaces* untuk membatasi blok. Python menyediakan *library* dengan standar yang besar yang dapat digunakan untuk berbagai aplikasi seperti pemrosesan bahasa alami (NLP), pembelajaran mesin (*machine learning*), analisis data, dll (Gupta & Negi, 2017).

Twitterscraper adalah *package* atau *library* dari *script* sederhana untuk mengambil data *tweet* menggunakan Python *package request* agar kontennya bisa diambil. Twitterscraper tidak ada limit atau batasan dalam pengambilan data, meskipun Twitter telah menyediakan REST API yang dapat digunakan oleh *developer* untuk mengakses dan membaca data Twitter. Tetapi salah satu kerugian terbesar dari Search API adalah orang-orang hanya dapat mengakses *tweet* yang

ditulis dalam 7 hari terakhir. Ini adalah hambatan utama bagi siapa saja yang mencari data masa lalu untuk membuat model (twitterscraper 0.2.7, 2016).

## 2.9. Penelitian Terdahulu

Tabel 2.1 menunjukkan beberapa jurnal yang digunakan sebagai tinjauan studi.

**Tabel 2. 1. Penelitian Terdahulu**

No	Judul	Pengarang	Kesimpulan	Perbandingan
1.	<i>Sentiment Analysis for Effective Stock Market Prediction</i>	Shri Bharathi, Angelina Geetha, B.S.Abdur Rahman (2017).	Kombinasi <i>sentiment polarity news sensex</i> points memberikan hasil yang efisien bagi peramal pasar saham kapan harus membeli atau menjual saham mereka. Hasil eksperimen kombinasi metode akurasi 78,75%. Meningkatkan 14,43% dibandingkan metode <i>moving average</i> saja yang sebesar 64,32% serta dengan algoritma standar ID3 dan C4.5.	Penelitian ini menggunakan <i>RSS Feed</i> berita perusahaan ARBK dan menggunakan <i>moving average stock level indicator</i> untuk menghitung <i>sensex points</i> . Sedangkan penelitian yang akan dilakukan menggunakan data berita tentang Facebook dari akun <i>news</i> di Twitter serta menggunakan <i>moving average</i> saja.

**Tabel 2. 1. Penelitian Terdahulu (Lanjutan)**

No	Judul	Pengarang	Kesimpulan	Perbandingan
2.	<i>Stock Trend Prediction Using News Sentiment Analysis</i>	Kalyani Joshi, Prof. Bharathi H. N., Prof. Jyothi Rao. K. J. Somaiya (2016).	Tren saham dapat diprediksi dengan artikel berita dan riwayat harga saham sebelumnya. Dibuktikan dari hasil membandingkan 3 model klasifikasi, yaitu <i>Random Forest</i> bekerja sangat baik dengan akurasi 86%-90%, nilai akurasi SVM sekitar 86% dan <i>Naïve Bayes</i> sekitar 83%.	Penelitian ini menggunakan data artikel berita sebagai sumber sentimen publik sedangkan penelitian yang akan dilakukan menggunakan data berita dari akun <i>news</i> di Twitter. Objek penelitian juga berbeda.
3.	<i>Stock Market Forecasting Based on Text Mining Technology: A Support Vector Machine Method</i>	Yancong Xie, Hongxun Jiang (2017) .	Metode SVM dengan teknologi <i>text mining</i> menunjukkan hasil yang sangat baik dalam memprediksi pasar saham, terutama harga saham yang spesifik. Media <i>online</i> ternyata cukup banyak memberikan pengaruh pada para investor untuk mengambil keputusan dalam transaksi pasar saham. Tetapi, metode SVM ini masih belum sempurna karena jika jumlah berita sedikit, hasil prediksinya tidak begitu baik.	Penelitian ini menggunakan data dari artikel berita finansial <i>online</i> di Cina sedangkan penelitian yang akan dilakukan mengambil data sentimen publik Twitter. Objek penelitian ini adalah pasar saham Cina dan penelitian selanjutnya tidak menggunakan model <i>classifier</i> .

U N I V E R S I T A S  
M U L T I M E D I A  
N U S A N T A R A