

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

#### 1.1.1 Identifikasi Masalah

Ilmu ekonomi merupakan salah satu ilmu terapan yang telah menggunakan teknik data mining sebagai alat prediksi dan peramalan. Karena dalam proses data mining, data di simpan secara elektronik dan dapat di cari atau ditambah secara otomatis oleh data itu sendiri. Hakekat data mining menurut [1]; adalah memecahkan masalah dengan menganalisis data yang sudah ada dalam database.

Contoh penggunaan data mining dalam ilmu ekonomi antara lain penentuan kelayakan kredit dengan menggunakan *credit scoring* [2]. Prediksi nilai tukar mata uang (*exchange rate*) dengan menggunakan *neural network* dan *genetic algorithm* (GA) [3]. Prediksi nilai saham dengan menggunakan kNN[4]. Peramalan harga saham dengan menggunakan data mining [5]. Peramalan pergerakan bursa saham menggunakan jaringan syaraf tiruan *neural network* [6].

Dari sekian banyak penelitian yang menggunakan data mining di bidang ekonomi, prediksi atas nilai saham merupakan objek yang sangat diminati. Saham memiliki definisi suatu nilai atau pembukuan dalam berbagai instrument finansial yang mengacu pada bagian kepemilikan perusahaan yang diterbitkan oleh perusahaan untuk memenuhi kebutuhan dana jangka panjangnya dalam bisnis dengan imbalan uang tunai [7].

Saham merupakan metode untuk meningkatkan modal perusahaan selain obligasi/ surat utang. Saham dapat berfluktuasi dalam jumlah dan nilai tidak seperti *asset/ harta* perusahaan lainnya. Bentuk dari variable transaksi harga saham biasanya terdiri dari kombinasi harga pembukaan, harga tertinggi, harga terendah, harga penutupan, serta volume penjualan saham.

Penelitian nilai saham dibidang bisnis yang telah dilakukan, secara garis besar hanya berfokus pada prediksi pergerakan harga saham. Penelitian tersebut didasarkan gerakan nilai saham di pasar saham pada masa yang lalu dan pada masa sekarang, kemudian nilai saham tersebut digunakan untuk memprediksi nilai saham di masa yang akan datang.

Padahal selain membeli dan menjual saham di pasar saham, masing-masing nilai saham tidak hanya ditandai oleh harga jual atau beli saham itu sendiri, tetapi juga oleh variabel lain. Faktor ekonomi yang berbeda-beda, stabilitas politik, dan kondisi yang tidak terduga lainnya adalah variabel yang telah dipertimbangkan untuk prediksi nilai saham merupakan contoh dari variabel-variabel lain yang mempengaruhi nilai saham [8].

Nilai saham memiliki karakteristik dinamis dan rentan terhadap perubahan cepat karena faktor internal yaitu sifat dasar dari domain keuangan dan sebagian besar karena campuran parameter yang dikenal antara lain (harga penutupan hari sebelumnya (*Previous Day's Closing Price*), rasio per *share* (*P/E Ratio*), dll.) dan factor yang tidak diketahui (hasil pemilu, rumor, dll) sebagai factor eksternal [9].

Penelitian prediksi nilai saham berdasarkan teknik yang digunakan dikelompokkan ke dalam 4 (empat) golongan pendekatan, yaitu; pendekatan teknik (*technical analysis*), pendekatan fundamental (*fundamental analysis*), pendekatan rentet waktu (*time series analysis*) dan pendekatan pembelajaran algoritma (*machine learning analysis*) [10].

Pendekatan teknis disebut sebagai pendekatan diagram (*chartist*). Karena menghasilkan prediksi berdasarkan nilai historis harga saham yang dipilih. Pada pendekatan fundamental, prediksi diketahui dari nilai sebenarnya saham dan membandingkannya dengan tingkat perdagangan saat ini dan merekomendasikan pembelian saham yang di perdagangan lebih rendah daripada nilai yang sebenarnya.

Pendekatan *time series* membagi model prediksi berdasarkan aliran linier yang dihasilkan dan nilai historis yang ditelusuri. Permodelan linier ini diklasifikasikan menjadi 2 (dua) jenis berdasarkan jumlah variabel yang terlibat dalam estimasi/prediksi tersebut.

Satu model variabel di sebut sebagai estimasi *univariate* dan model dan beberapa nomor dari variabel yang di pertimbangkan dalam model disebut sebagai model estimasi regresi *multivariate*. Terakhir, prediksi dengan menggunakan pembelajaran algoritma atau *machine learning analysis* merupakan metode yang mempertimbangkan *sample* dataset untuk menghasilkan pola linier dan nonlinier dari data.

### **1.1.2 Analisis Masalah**

Pasar saham atau bursa saham serupa dengan permainan menang kalah/ *game winners and loser* di bidang keuangan [11]. Nilai saham dapat berubah sewaktu-waktu, ada faktor penentu naik turunnya nilai saham. Sehingga penelitian tentang cara yang benar-benar tepat untuk mengetahui variabel mempengaruhi pasar dan perusahaan yang jumlahnya tidak terhitung menjadi fokus di bidang ekonomi dan finansial.

Prediksi nilai saham yang tepat, memberikan kontribusi yang besar untuk menentukan area portofolio manajemen, area pengalokasian *asset* atau harta perusahaan dan terlebih pada meminimalisasi resiko manajemen [12]. Ada empat macam pendekatan untuk prediksi nilai saham, antara lain pendekatan secara teknis, pendekatan secara fundamental, pendekatan secara rentet waktu dan pendekatan dengan menggunakan algoritma pembelajaran *machine learning*.

Pendekatan teknis disebut sebagai pendekatan diagram (*chartist*) karena metode ini mengevaluasi efek dengan menganalisis statistik yang dihasilkan oleh aktivitas pasar, seperti harga terakhir dan volume. Analisis teknikal tidak mencoba untuk mengukur nilai intrinsik suatu sekuritas, namun menggunakan grafik dan

alat-alat lain untuk mengidentifikasi pola-pola yang dapat menyarankan aktivitas masa depan. Beberapa metode prediksi saham dengan pendekatan teknis antara lain; *Variable Moving Average* (VMA), *Trading Range Breakout* (TRB), *Bollinger Bands*, *Money Flow Index* (MFI).

Pada pendekatan fundamental, prediksi di ketahui dari nilai sebenarnya saham dan membandingkannya dengan tingkat perdagangan saat ini dan merekomendasikan pembelian saham yang di perdagangan lebih rendah daripada nilai yang sebenarnya. Metode yang digunakan pada pendekatan fundamental antara lain; *Consumer Price Index* (CPI), *Consumer Sentiment Index*, *Consumer Confidence Index*, *New Orders Diffusion Index*, *Leading Indicators Index*, *Interest Rate of the 30-Year Conventional Mortgage*, *Federal Funds Rate*, *Price-Earnings* (P/E) *Ratio*.

Beberapa penelitian dengan menggunakan pendekatan fundamental dilakukan antara lain oleh Jaouida Elleuch [13]. Penelitian ini meletakkan dasar-dasar pengukuran akuntansi sebagai nilai standar untuk memprediksi nilai saham di masa yang akan datang. Penelitian ini menghasilkan nilai prediksi yang akurat karena menggunakan dasar-dasar pengukuran akuntansi antara lain;

1. Persediaan/ *Inventory* (INV);
2. Piutang Usaha/ *Accounts Receivable* (AR);
3. Investasi/ *Investments* (INVES);
4. Margin Kotor/ *Gross Margin* (GM);
5. Tenaga Kerja/ *Labour Force* (LF);
6. Laba atas harta dan variasi laba atas harta/ *Return on assets* (ROA);
7. *Variation in Return On Assets* ( $\Delta$ ROA);
8. Aliran kas/ *Cash flow* (CF);
9. Pendapatan atau Biaya yang masih harus dibayar/ *Accruals* (ACC);
10. *Leverage*  $\Delta$ LEV;
11. Likuiditas/ *Liquidity* ( $\Delta$ LIQUID);
12. Perputaran harta/ *Assets Turnover* ( $\Delta$ TURN).

Akan tetapi, pada pendekatan fundamental ini memiliki kelemahan dengan banyaknya waktu yang di butuhkan pengerjaan perhitungan analisis.

Pendekatan *time series* membagi model prediksi berdasarkan aliran linier yang dihasilkan dan nilai historis yang ditelusuri. Permodelan linier ini diklasifikasikan menjadi 2 (dua) jenis berdasarkan jumlah variabel yang terlibat dalam estimasi/prediksi tersebut. Satu model variabel di sebut sebagai estimasi *univariate* dan model dan beberapa nomor dari variabel yang di pertimbangkan dalam model disebut sebagai model estimasi regresi *multivariate*.

Diambil dari Takashi Shibuya. [14] Prediksi *time series* merupakan isu penting dalam berbagai berbagai bidang seperti keuangan [4], prakiraan cuaca [1], dan perencanaan transportasi [7]. Sebagai contoh, ekonom memprediksi tren ekonomi dari indikator seperti pertukaran tarif, harga saham, dan PDB. Kesimpulan yang ditarik oleh para ekonom kemudian mempengaruhi keputusan *trader*. Tindakan pedagang kemudian dipantulkan kembali ke dalam indikator ekonomi, sehingga menciptakan siklus yang tidak berujung. Berdasarkan jumlah variabel yang terlibat dalam estimasi/ prediksi Prediksi nilai saham dengan menggunakan pendekatan *time series* di bagi menjadi 2 yaitu prediksi *time series univariate* atau satu model variable dan *prediksi time series multivariate* dengan mempertimbangkan beberapa nomor *variabel* yang mempengaruhi prediksi.

Pendekatan *timeseries* dengan *univariate* untuk prediksi nilai saham pernah dilakukan oleh Illka Virtanen *et al*. Metode yang di gunakan adalah ARIMA (*Auto Regressive Moving Average*) dengan metode *Box Jenkins*. Satu variabel harga yang digunakan di bagi menjadi harga bulanan, kuartal dan harian.

Prediksi nilai saham dengan menggunakan *multivariate* pernah dilakukan oleh Alessia Naccarato *et al* [15]. Model yang digunakan adalah bivariat berko-integrasi vektor *autoregressive*. Nilai intrinsik saham sebagai variabel yang digunakan

untuk pemilihan portofolio saham. Nilai instrinsik itu antara lain dasar perkiraan laba rata-rata dari saham dan volatilitas harga saham dari 150 saham yang diperdagangkan di Pasar saham Italia. April Kerby *et al* [11], juga melakukan penelitian prediksi nilai saham dengan *multivariate* dengan tujuan menentukan nilai saham yang baik untuk digunakan sebagai investasi ataukah tidak.

Pendekatan *machine learning analysis* pada data nilai saham di bursa saham melakukan prediksi dengan menerapkan kombinasi dari analisis teknis (*technical analysis*) dan pendekatan fundamental (*fundamental analysis*). Secara teknis, *machine learning* bertujuan melihat apakah algoritma yang digunakan cukup akurat untuk mempelajari pola yang mendasari dalam data rentet waktu *time series* dari saham dan *fundamental analysis* memainkan peran utama dalam mengevaluasi dan peramalan kinerja perusahaan dan parameter lain yang serupa. Penelitian dengan menggunakan *machine learning* antara lain dilakukan oleh V. H Shah *et al*, dengan menggunakan metode *Decision Stump*, *Linier Regression*, *Support Vector Machine* dan algoritma *Adaabost* [16]. Penelitian ini menghasilkan teori bahwa harga saham tidak terpengaruhi oleh harga saham sebelumnya (historis), sehingga tidak ada *trend* dari nilai saham. Prediksi nilai saham dengan GA (*Genetic Algorithm*) dilakukan oleh R. L. Naik, D. Ramesh, B. Manjula, and A. Govardhan, di tahun 2012 dengan mengambil data saham dari Bursa Saham Semen India (*India Cement Stocks Index*), menunjukkan hasil GA memberikan nilai prediksi indeks tertinggi di bandingkan dengan metode yang lain sehingga sangat akurat untuk menentukan keputusan jual atau beli saham, akan tetapi memiliki kelemahan pada hanya dapat digunakan untuk periode waktu prediksi singkat, karena hanya berdasarkan nilai historis [17].

Penggunaan Algoritma k-NN (*k-Nearest Neighbor*) digunakan untuk prediksi nilai saham yang di terapkan pada harga penutupan saham di tahun 2013 oleh M. K. A. Shatnawi, hasil prediksi menunjukkan bahwa nilai prediksi lebih besar dari yang sebenarnya.

Prediksi nilai saham dengan menggunakan algoritma *Artificial Neural Network* (ANN) telah banyak dilakukan, antara lain oleh N. Masoud, di tahun 2014 [18]; menunjukkan bahwa ANN merupakan algoritma yang baik untuk memprediksi nilai harga saham harian. Di tahun 2008, algoritma ANN ini digunakan untuk memprediksi nilai saham oleh R. A. Y and S. Price [19], dengan hasil ANN atau JST (Jaringan Syaraf Tiruan) merupakan alat yang berguna dalam memprediksi pergerakan harga saham tertentu. Dalam jangka pendek, hubungan harga antara unsur-unsur dari sektor sangat tegas dan jelas.

Sebuah JST dapat belajar hubungan harga dengan tingkat akurasi yang tinggi dan digunakan untuk menghasilkan keuntungan. Dengan jumlah data yang cukup besar, sebaiknya pada saat *volatilitas* atau pergerakan saham rendah dan selama periode waktu yang singkat, JST meniru jenis indeks arbitrase statistik sehingga dapat digunakan untuk pada saham apapun.

Dari metode – metode *machine learning* yang pernah digunakan, pasti memiliki kelebihan dan kelemahan. *Support Vector Machine* (SVM) memiliki kelebihan dapat meminimalkan resiko [20] dan dapat mengatasi permasalahan pada *neural network* yang terjebak pada *local minima* [21], namun SVM memiliki kelemahan pada kesulitan pemilihan parameter, karena tidak ada keseragaman dalam kriteria pemilihan parameter [22].

Keistimewaan ANN dalam masalah prediksi, diantaranya adalah ANN merupakan metode yang adaptif, karena memiliki kemampuan generalisasi jaringan yang tetap kuat dan akurat dalam *dataset* yang dapat berubah dari waktu ke waktu [23].

Keistimewaan lainnya adalah setelah melalui proses *training*, ANN memiliki kinerja yang baik bila diterapkan pada data diluar sampel [24], meskipun data yang didapatkan kurang lengkap dan rusak [25]. Selain itu, jaringan pada ANN juga dapat memprediksi pada kelas yang besar dengan tingkat akurasi yang tinggi [25]. Namun, ANN juga memiliki kelemahan, diantaranya adalah ketergantungan

pada beberapa parameter, seperti *network size*, *learning parameters* dan pemilihan *initial weight*, serta *neural network* memiliki konvergensi yang lambat karena terjebak pada *local minima* [20].

*Backpropagation* merupakan algoritma yang sering dipakai untuk melakukan proses *learning* pada *neural network*. Algoritma ini banyak digunakan untuk memecahkan banyak masalah di dunia nyata dengan membangun model terlatih dan telah menunjukkan kinerja yang baik dalam beberapa masalah *non linier* [26]. Namun, *backpropagation* memiliki kelemahan, diantaranya adalah *backpropagation* terjebak pada *local minima* dan konvergensi yang lambat sehingga menyebabkan *backpropagation* sangat tergantung pada parameter awal, seperti *input nodes*, *hidden nodes*, *output*, *learning rate* dan bobot koneksi dalam jaringan [26].

Secara umum, memilih struktur jaringan pada ANN yang sesuai sangat penting dan sulit, karena jaringan yang lebih kecil dari yang dibutuhkan mungkin tidak dapat memberikan kinerja yang baik karena terbatasnya informasi pada proses *training*, sedangkan jaringan yang lebih besar dari yang dibutuhkan kemungkinan juga memiliki hubungan yang berlebihan diantara jaringan-jaringan yang terbentuk [27].

Penelitian ini digunakan untuk melakukan prediksi nilai saham dengan melakukan komparasi hasil prediksi menggunakan ANN terhadap faktor-faktor yang mempengaruhi nilai saham. Prediksi dilakukan dengan menentukan klasifikasi faktor penentu nilai saham menjadi faktor *univariate* atau *multivariate* terlebih dahulu.

Berdasarkan hal tersebut diatas, maka dalam penelitian ini pendekatan yang tepat digunakan dalam prediksi nilai saham adalah pendekatan rentet waktu (*time series*), dengan membandingkan model variabel yang di pertimbangkan dalam proses prediksi yaitu *univariate* dan *multivariate*.

Dataset penelitian yang digunakan pada penelitian mengambil dari web *Yahoo Finance* yaitu *Tata Consultancy Service (TCS)* dan *Indonesia Exchange Rate (IDX)*. Tujuan penelitian ini untuk menemukan nilai prediksi nilai saham yang mendekati nilai *real/* nyata. Nilai prediksi di bandingkan antara prediksi dengan menggunakan satu variabel (*univariate*) dan beberapa variabel (*multivariate*).

Perbandingan atau komparasi tersebut di ukur dengan tingkat *error* prediksi berdasarkan pada variabel-variabel yang mempengaruhi. Dengan asumsi bahwa selalu ada beberapa dependensi atau ketergantungan antara bidang yang berbeda di pasar saham.

Ketergantungan antar variabel di olah dengan menggunakan algoritma ANN. Tujuannya adalah untuk menemukan apakah setelah dilakukan komparasi terhadap variabel- variabel dari perusahaan skala besar dengan asumsi adanya ketergantungan antara saham perusahaan besar mempengaruhi saham perusahaan skala kecil atau sebaliknya. Algoritma ini dianggap cocok untuk permodelan data non parametrik dalam berbagai konteks yang berbeda di mana output adalah fungsi non-linear dari input. Oleh karena itu, penelitian ini akan melakukan komparasi *univariate* dan *multivariate* dengan menggunakan ANN untuk prediksi nilai saham.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang masalah yang telah diuraikan diatas, dapat disimpulkan bahwa rumusan masalah pada penelitian ini adalah:

1. Belum diketahui nilai prediksi yang terbaik pada dataset nilai saham *univariate* dan *multivariate* yang mempengaruhi nilai saham dengan penerapan ANN.

### 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, maka tujuan dari penelitian ini adalah:

1. Menentukan nilai prediksi yang terbaik pada dataset nilai saham dengan menggunakan *Root Mean Square Error* (RMSE) terkecil perbandingan antara *univariate* dan *multivariate*.

### 1.4 Manfaat Penelitian

1. Hasil penelitian ini diharapkan dapat digunakan oleh investor, pemerintah dan masyarakat sebagai masukan untuk penentuan kebijakan di bidang investasi saham dengan lebih mudah, cepat dan akurat.
2. Hasil penelitian ini diharapkan dapat bermanfaat untuk masyarakat yang memiliki keterkaitan dengan perubahan nilai saham, seperti para pelaku bursa saham.
3. Hasil penelitian ini diharapkan dapat menyempurnakan metode yang pernah dilakukan dalam penerapan *neural network* dalam melakukan prediksi nilai saham.

### 1.5 Sistematika Penulisan

Penulisan pada penelitian ini akan dibagi menjadi enam bagian, yaitu:

#### BAB I Pendahuluan

Bab ini membahas tentang latar belakang penelitian tentang prediksi nilai saham, rumusan permasalahan yang ditemukan, tujuan penelitian, manfaat penelitian, dan sistematika penulisan tesis.

#### BAB II Tinjauan Pustaka

Bab ini berisi tentang penelitian-penelitian terkait khususnya mengenai *state-of-the-art* dari penelitian tentang prediksi nilai saham dan tinjauan pustaka untuk teori-teori yang digunakan.

Dalam bab ini juga diuraikan tentang kerangka pemikiran yang merupakan penjelasan tentang kerangka berpikir untuk memecahkan masalah yang sedang diteliti, termasuk menguraikan objek penelitian.

### BAB III Metode Penelitian

Bab ini akan menjelaskan metode penelitian yang digunakan, secara umum terdiri dari teknik pengumpulan data yang digunakan, proses pengolahan awal data, eksperimen dan pengujian metode, serta evaluasi dan validasi hasil.

### BAB IV Tahapan Penelitian

Bab ini berisi tahapan dari penelitian yang dilakukan mulai dari pengolahan data, proses pembagian data menjadi *univariate* dan *multivariate*.

### BAB V Hasil Penelitian dan Pembahasan

Bab ini akan berisi pembahasan dari hasil eksperimen yang dilakukan serta evaluasi dan validasi. Hasil eksperimen akan berisi data yang disajikan dalam bentuk tabel-tabel dan hasil analisa tingkat akurasi dari model yang digunakan.

### BAB VI Penutup

Bab ini berisi kesimpulan dari hasil penelitian dan saran dari penelitian yang telah dilakukan.