

# **Analisis Harga Saham PT Astra Internasional Tbk Menggunakan Data Dari Bursa Efek Indonesia dalam Jangka Waktu Pendek Menggunakan Metode Naïve Bayes dan Decision Tree-J48**

**Heima Sitorus dan Yusran Tarihoran  
Fakultas Teknologi Informasi, Universitas Advent Indonesia**

## **Abstrak**

Prediksi harga saham merupakan tindakan yang sangat diperlukan sebagai langkah awal untuk mengambil keputusan untuk menentukan kapan akan melakukan transaksi penjualan dan pembelian sebuah saham yang ada di bursa efek Indonesia. Prediksi harga saham dilakukan dengan menggunakan dua metode yang akan dibandingkan, yaitu; Metode Naïve Bayes dan Metode *Decision Tree*-J48. Jumlah data yang digunakan adalah sebanyak 1.195 dan data yang digunakan sebagai data *testing* adalah sebanyak 20% atau 239 data. Adapun hasil dari tingkat akurasi dari Metode Naïve Bayes menggunakan data *testing* adalah 92.0502 %. Dengan presentasi untuk nilai *precision* atau tingkat ketepatan informasi yang diharapkan oleh penulis dengan jawaban yang diberikan oleh sistem sebesar 0.920 dan nilai untuk *recall* atau tingkat keberhasilan terhadap informasi yang ditemukan kembali oleh sistem sebesar 0,961.

Sedangkan hasil prediksi dengan menggunakan metode *Decision Tree* J-48 untuk hasil akurasi dengan menggunakan data *testing* adalah sebesar 98.7448 %. Dengan presentasi untuk nilai *precision* atau tingkat ketepatan informasi yang diharapkan oleh penulis dengan jawaban yang diberikan oleh sistem. sebesar 0.989 dan nilai untuk *recall* atau tingkat keberhasilan terhadap informasi yang ditemukan kembali oleh sistem sebesar 0.997

**Kata Kunci:** *Harga saham, Bursa Efek Indonesia, Naïve Bayes, Decision tree-J48, WEKA.*

## ***Analysis of PT Astra International Tbk's Share Price Using Data from the Indonesia Stock Exchange in Short Term Period Using the Naïve Bayes Method and Decision Tree-J48***

### ***Abstract***

*Stock price prediction is a very necessary action as a first step to taking the decision to determine when to transact a sale and purchase of existing shares in Indonesia stock exchange. Stock price prediction is done using two methods are compared, i.e.; Naïve Bayes method and the method of Decision Tree-J48. The amount of data used is as much as 1,195 and data used as the data for testing is as much as 20% or 239 the data.*

*As a result of the level of accuracy of the Naïve Bayes Method using data testing is 92.0502%. With the presentation for the value of precision or accuracy of information expected by the author with the answers given by the system of 0920 and the value for the recall or the success rate of taking action against information found by the system of 0.961.*

*While the results of a prediction using Decision Tree J-48 results for accuracy using data testing is 98.7448%. %. With the presentation for the value of precision or accuracy of information expected by the author with the answers given by the system. of 0.989 and value for the recall or the success rate of taking action against information found by the system of 0.997.*

**Keywords:** *Stock price, Indonesia stock exchange, Naïve Bayes, Decision tree-J48, WEKA.*

## **Pendahuluan**

Pasar modal dapat mempermudah masyarakat dan para investor untuk mengetahui perkembangan tentang harga saham setiap perusahaan-perusahaan di Indonesia yang sudah *go public*. Salah satu perusahaan yang sudah terdaftar di bursa efek Indonesia dan telah *go public* adalah PT Astra Internasional Tbk. Masyarakat yang ingin mengetahui harga saham PT Astra Internasional Tbk setiap harinya dapat melihatnya di bursa efek Indonesia.

Keterbukaan data harga saham yang dimiliki oleh PT Astra Internasional Tbk, hal ini dapat digunakan oleh masyarakat dan para investor untuk mendapatkan keuntungan. Namun, untuk mendapatkan keuntungan, maka masyarakat dan para investor harus mengetahui waktu yang tepat untuk melakukan penjualan dan pembelian saham yang diinginkan di PT Astra Internasional Tbk.

## Landasan Teori

### Harga Saham

Harga saham adalah harga penutup yang sering ditampilkan di bursa efek Indonesia atau di pasar modal Indonesia yang menampilkan grafik atau harga yang dapat digunakan masyarakat, investor dan perusahaan-perusahaan yang ingin melakukan penjualan dan pembelian saham, sehingga memberikan keuntungan secara pribadi.

Menurut Supranto. J menyatakan bahwa "Indeks harga saham adalah angka indeks harga saham yang telah disusun dan dihitung sedemikian rupa sehingga menghasilkan tren. Angka indeks merupakan angka yang dibuat sedemikian rupa sehingga dapat dipergunakan untuk membandingkan kegiatan ekonomi atau peristiwa, bisa berupa perubahan harga saham dari waktu ke waktu".

Untuk dapat melakukan analisis harga saham, ada dua teknik analisis yang sering digunakan, yaitu:

1. Analisis Fundamental

Analisis fundamental adalah metode analisis yang didasari pada ekonomi suatu perusahaan. Analisis ini juga sering digunakan untuk melakukan analisis perusahaan dalam bidang keuangan. Analisis ini juga sering digunakan oleh pihak-pihak yang ingin mengetahui tentang harga saham dalam jangka waktu yang lama atau dalam jangka waktu yang panjang.[2]

2. Analisis Teknikal

Analisis teknikal adalah analisis yang sering menggunakan data-data yang sudah lampau, terutama pergerakan harga dan *volume*. Analisis teknikal merupakan suatu metodologi untuk memprediksi harga saham berdasarkan kondisi penawaran dan permintaan terhadap saham tersebut.[2] Untuk dapat menganalisis harga saham, maka hal yang akan diperhatikan penulis adalah sebagai berikut:

- *Prev* atau *previous* adalah harga penutupan saham pada hari bursa sebelumnya. Misalnya jika hari ini adalah hari Selasa, maka *prev* menunjukkan harga penutupan pada hari sebelumnya, yaitu hari Senin.
- Harga tertinggi adalah harga tertinggi saham yang terjadi hari itu juga
- Harga terendah adalah harga yang paling rendah yang terjadi hari itu.
- Frekuensi perdagangan saham adalah berapa kali transaksi jual beli terjadi pada saham yang bersangkutan pada waktu tertentu.
- Individual indek merupakan jumlah indeks saham yang ada di dalam pasar modal.
- *Volume* ini menentukan jumlah transaksi yang diperdagangkan dalam pasar pada satu periode tertentu.
- *Value* merupakan jumlah saham yang terjual dalam satu waktu.

## Metode Penelitian

### Data Mining

Data *Mining* (DM) adalah salah satu bidang yang saat ini mengalami perkembangan yang sangat pesat karena besarnya kebutuhan terhadap nilai tambah dari *database* berskala besar, yang makin banyak terakumulasi sejalan dengan pertumbuhan teknologi informasi. Definisi umum dari Data *mining* itu sendiri merupakan proses yang dilakukan untuk menggali nilai tambah berupa pengetahuan dari sekumpulan data yang selama ini tidak dianggap penting oleh masyarakatnya.

Data *mining*, sering juga disebut sebagai *Knowledge Discovery in Database* (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar.

Menurut Achmad Basuki Proses dalam tahapan data mining terdiri dari tiga langkah utama, yaitu:

1. Data *Preparation*. Pada langkah ini data dipilih dibersihkan dan dilakukan *preprocessed*.
2. Penggunaan Algoritma data *mining*.
3. Tahap Analisa Keluaran dari data *mining* dievaluasi untuk melihat apakah *knowledge* domain ditemukan dalam bentuk *rule* yang telah diekstrak.



**Gambar 1** Langkah-langkah dalam Proses Data Mining

### Algoritma Naive Bayes

Algoritma *Naive Bayes* merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. *Naive Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris. *Thomas Bayes*, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai *Teorema Bayes*. Teorema tersebut dikombinasikan dengan *Naive* dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi *Naive Bayes* diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya. Persamaan dari teorema *Bayes* adalah:

$$\begin{aligned}
 P(s|x) &= \frac{\operatorname{argmax}_{x \in X} P(y|x) P(x)}{P(X)} \\
 &= \frac{\operatorname{argmax}_{x \in X} P((y|x) P(X)}{P(X)}
 \end{aligned}$$

### Metode Decision Tree J-48

Metode Decision Tree J-48 merupakan salah satu kelas yang terdapat pada program data mining yang ada di *software* weka. Algoritma J48 merupakan implementasi dari program algoritma C4.5

Maka secara umum Algoritma J-48 biasanya digunakan untuk masalah klasifikasi. Sebuah decision tree terdiri dari beberapa simpul yaitu *tree's root*, internal *node* dan *leafs*. Konsep entropi digunakan untuk penentuan pada atribut mana sebuah pohon terbagi (*split*). Semakin tinggi *entropy* sebuah sampel, semakin tidak murni sampel tersebut.

Rumus yang digunakan untuk menghitung "*entropy*" sampel S adalah;

$$Entropy(S) = \sum \tau - p_i \log_2 p_i$$

Atau

$$Entropy(S) = -p_1 \log_2 p_1 - p_2 \log_2 p_2$$



**Gambar 2** Konsep Decision Tree[4]

### Waikato Environment Knowledge And Analysis (WEKA)

WEKA adalah sebuah perangkat lunak data mining/machine learning yang dibangun oleh *Department of Computer Science University of Waikato* di *New Zealand*. WEKA mampu untuk menyelesaikan masalah-masalah yang berhubungan dengan data mining yang ada didunia nyata, khususnya klasifikasi yang mendasari pendekatan terhadap machine learning.

Secara umum WEKA menyediakan 3 fasilitas untuk data mining. Ketiga fasilitas tersebut adalah data processing, data mining dan visualization tools.

### Analisa dan Metode Penelitian

Pengumpulan data dan Pengolahan Data Saham BEI. Data penelitian yang digunakan oleh penulis adalah data yang telah di publikasikan melalui *website* Bursa Efek Indonesia (<http://www.idx.co.id/id-id/beranda/unduhdata/ringkasan.aspx>.)

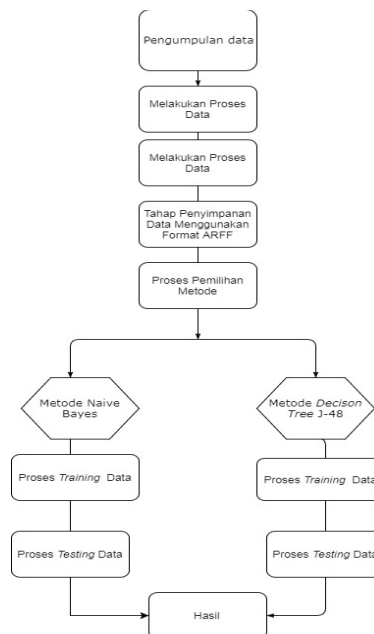
### Data Saham

Pengumpulan data PT Astra Internasional Tbk yang dilakukan secara manual. Data yang diambil penulis adalah data harian yang mulai dari tanggal 1 Januari 2013 sampai dengan 30 November 2017. Namun data yang telah didapatkan dari *website* resmi bursa efek Indonesia masih harus dikelolah ulang, karena data yang dihasilkan dari *website* tersebut masih merupakan gabungan dari data harga saham perusahaan-perusahaan yang telah terdaftar.

NO	Date	Previous	High	Low	Close	Volume	Value	Frequency	Output
1	2-Jan-13	7600	7700	7450	7500	23374000	1.77427E+11	1753	Naik
2	3-Jan-13	7500	7900	7500	7850	70205000	5.42508E+11	4567	Turun
3	4-Jan-13	7850	7850	7750	7850	29485000	2.30094E+11	2989	Turun
4	7-Jan-13	7850	7850	7700	7750	20890000	1.62279E+11	1863	Turun
5	8-Jan-13	7750	7800	7600	7650	24269500	1.86656E+11	1969	Turun
6	9-Jan-13	7650	7750	7650	7700	38907000	2.99253E+11	1936	Naik
7	10-Jan-13	7700	7700	7300	7350	50290500	3.74987E+11	5175	Turun
8	11-Jan-13	7350	7500	7300	7300	33422500	2.46395E+11	3791	Naik
9	12-Jan-13	7300	7500	7350	7500	28014000	2.08141E+11	2429	Turun
10	13-Jan-13	7500	7600	7500	7600	32451000	2.44975E+11	2845	Naik
11	16-Jan-13	7600	7650	7500	7650	23595000	1.79349E+11	2810	Turun
12	17-Jan-13	7650	7650	7500	7500	27764000	2.1031E+11	2403	Turun
13	18-Jan-13	7500	7750	7550	7750	34305000	2.63287E+11	2963	Naik

Gambar 3 Contoh Daftar harga saham PT Astra Internasional Tbk 2013-2017

### Block Diagram Penelitian



Gambar 4 Block Diagram Penelitian

Keterangan *Block Diagram*:

1. Pengumpulan data merupakan proses pengambilan data yang terdapat dari *website* PT Astra Internasional Tbk dengan periode 1 Januari 2013 sampai dengan periode 30 November 2017.
2. Melakukan proses data yang pertama adalah memisahkan data PT Astra Internasional Tbk kedalam satu file dan disimpan di Microsoft Excel 2013.
3. Proses data yang kedua adalah menyimpan data dengan format.Cvs.
4. Setelah data disimpan dengan format.Cvs, bukalah *software* WEKA, dan bukalah file, dan simpan dengan format.Arff.
5. Lalu, pilih tool *Classify*, dan pilihlah metode yang diinginkan. Dalam penelitian ini penulis menggunakan metode Naïve Bayes dan *Decision Tree* J-48.
6. Setelah memilih metode yang digunakan, maka langkah selanjutnya yang dilakukan adalah membuka file yang telah disimpan dengan format.Arff, maka pilihlah *use training set* untuk mendapatkan hasil data *training* dengan menggunakan metode Naïve Bayes dan *Decision Tree* J-48.
7. Hasil data *training* yang sudah didapat, maka harus dijadikan menjadi data acuan untuk mendapatkan hasil data *testing*. Untuk mendapatkan hasil data *testing*, pilihlah data yang dijadikan diuji sebagai data *testing*, lalu pilihlah tool *supplied test set*, dan pilih data yang diuji oleh WEKA.
8. Setelah melakukan proses *training* dan proses *testing* maka akan mendapatkan hasil data *training* data *testing* dengan menggunakan metode Naïve Bayes dan *Decision Tree* J-48.

## Hasil Pengujian Dan Penelitian

### Pengujian Data

Pengujian data dilakukan dengan menggunakan bantuan *software* yang dapat sekaligus digunakan untuk pengolahan data, yaitu adalah WEKA. Data harga saham yang telah dikumpulkan dan disimpan di program *Excel*, lalu disimpan dengan format CVS (*Concurrent Versioning System*).

Selanjutnya data yang telah disimpan dapat dibuka di program weka, selanjutnya disimpan ulang dengan format ARFF (*Andrew's Ridiculos File Format*).

### Pengujian Data *Training*

Jumlah data yang akan di *training* yaitu sebanyak 1195 yang Merupakan Data yang sudah disiapkan dan yang sudah berbentuk Arff.

**Tabel 1** Hasil Naïve Bayes dan *Decision Tree* J-48 di WEKA

Jumlah Saham	Harga	Naïve Bayes	<i>Decision Tree</i> J-48
Benar		1010	1182
Salah		185	13

Keterangan Tabel 1 yaitu:

1. Naïve Bayes  
Terdapat 1010 data harga saham yang ditentukan oleh WEKA yang diklasifikasikan benar. Dan sebanyak 185 data harga saham yang dinyatakan oleh WEKA yang diklasifikasikan mengalami kesalahan.
2. *Decision Tree* J-48  
Terdapat 1182 data harga saham yang ditentukan oleh WEKA yang diklasifikasikan benar. Dan sebanyak 13 data harga saham yang dinyatakan oleh WEKA yang diklasifikasikan mengalami kesalahan.

```

=== Summary ===
Correctly Classified Instances      1010      84.5188 %
Incorrectly Classified Instances    185       15.4812 %
Kappa statistic                    0.5551
Mean absolute error                 0.1717
Root mean squared error             0.3083
Relative absolute error             44.3925 %
Root relative squared error         70.1218 %
Total Number of Instances          1195
    
```

**Gambar 5** Summary dengan menggunakan metode Naïve Bayes

Keterangan berdasarkan Gambar 5 adalah sabagai berikut:

1. *Correctly Classified Instance*

Rumus untuk mencari nilai akurasi adalah:

$$\text{Akurasi} = \frac{\text{jumlah prediksi harga saham yang benar}}{\text{Total jumlah data prediksi}}$$

$$\text{Akurasi} = \frac{f_0 + f_3}{f_0 + f_1 + f_2 + f_3}$$

$$\text{Akurasi} = \frac{838 + 172}{838 + 141 + 44 + 172}$$

$$\text{Akurasi} = \frac{1010}{1195} = 0.845188$$

$$\text{Presentasi Akurasi} = 0.8451 * 100 = 84.5188\%$$

2. *Incorrectly Classified Instance*

$$\text{Error Rate} = \frac{\text{jumlah prediksi harga saham yang salah}}{\text{Total jumlah data prediksi}}$$

$$\text{Error Rate} = \frac{f_1 + f_2}{f_0 + f_1 + f_2 + f_3}$$

$$\text{Error Rate} = \frac{141+44}{838+141+44+172} =$$

$$\text{Akurasi} = \frac{185}{1195} = 0.154812$$

$$\text{Presentasi Akurasi} = 0.154811 * 100 = 15.4812\%$$

3. *Total Number Of Instance*

*Total Number of Instance* merupakan jumlah data yang digunakan dalam proses *training* yaitu sebanyak 1.195.

```

=== Confusion Matrix ===

  a  b  <-- classified as
838 44 |   a = Naik
141 172 |   b = Turun
    
```

**Gambar 6** Confusion Matrix

Keterangan dari Gambar 6 adalah sebagai berikut:

Pada baris pertama menerangkan bahwa terdapat sebanyak 838 data harga saham dari *classified* "Naik" yang diklasifikasikan mengalami kenaikan harga, dan terdapat 44 data harga saham yang seharusnya mengalami kenaikan harga saham, namun salah diklasifikasikan menjadi mengalami penurunan harga saham.

Pada baris kedua menerangkan bahwa terdapat 172 data harga saham dari *classified* "Turun" yang diklasifikasikan mengalami penurunan harga saham, namun terdapat 141 data harga saham yang seharusnya mengalami penurunan harga saham, namun yang salah diklasifikasikan menjadi mengalami kenaikan harga saham.

```

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
      0.950   0.450   0.856   0.950   0.901   0.571  0.940   0.975   Naik
      0.550   0.050   0.796   0.550   0.650   0.571  0.940   0.858   Turun
Weighted Avg.  0.845   0.346   0.840   0.845   0.835   0.571  0.940   0.944
    
```

**Gambar 7** Detailed Accuracy

Keterangan Gambar 7 adalah:

1. *True positif Rate (TP Rate)*

*TP Rate* atau *True positif Rate* adalah jumlah tingkat kebenaran dari data harga saham yang diklasifikasikan sebagai kelas x, diantara semua data harga saham yang benar memiliki kelas x yang bernilai sama dengan *recall*.

Untuk mencari nilai TP Rate pada kelas harga saham yang *Naik* adalah:

$$\begin{aligned}
 \text{True Positif Rate} &= \frac{\text{Jumlah data yang benar}}{\text{Jumlah data yang benar} + \text{Jumlah data yang salah}} \\
 \text{True Positif Rate} &= \frac{838}{838 + 44} = 0,950
 \end{aligned}$$

Untuk mencari nilai FN atau *False Negative* pada kelas harga saham yang *Turun* adalah:

$$\begin{aligned}
 \text{False Negative} &= \frac{\text{Jumlah data yang salah}}{\text{Jumlah data yang benar} + \text{Jumlah data yang salah}} \\
 \text{True Positif Rate} &= \frac{172}{172 + 141} = 0,550
 \end{aligned}$$

2. *False Positive (FP Rate)*

Untuk mencari nilai *False Positive* adalah nilai harga saham yang diklasifikasikan sebagai kelas x, tetapi masuk kedalam kelas yang berbeda, dan yang bukan dari kelas x itu sendiri:

Untuk mencari nilai *False Positive* pada kelas yang memiliki data harga saham naik:

$$False\ Positive = \frac{Jumlah\ data\ yang\ benar}{Jumlah\ data\ yang\ benar + Jumlah\ data\ yang\ salah}$$

$$False\ Positive = \frac{141}{141 + 172} = 0,450$$

Untuk mencari *False Negative* pada kelas yang turun adalah:

$$False\ Negative = \frac{Jumlah\ data}{Jumlah\ data\ yang\ benar * Jumlah\ data\ yang\ salah}$$

$$False\ Negative = \frac{838 + 44 + 141 + 172}{141 * 172} = 0,050$$

### 3. *Precision*

*Precision* merupakan tingkat ketepatan informasi yang diharapkan oleh penulis dengan jawaban yang diberikan oleh sistem.

Untuk mencari nilai *Precision* dari kelas Naik:

$$Precision = \frac{Jumlah\ data\ benar}{Jumlah\ data\ yang\ benar + Jumlah\ data\ yang\ salah}$$

$$Precision = \frac{838}{838 + 141} = 0,856$$

Untuk mencari nilai *Precision* dari kelas Turun:

$$Precision = \frac{Jumlah\ data\ salah}{Jumlah\ data\ yang\ benar + Jumlah\ data\ yang\ salah}$$

$$Precision = \frac{172}{172 + 44} = 0,796$$

### 4. *Recall*

*Recall* merupakan tingkat keberhasilan sistem dalam menemukan kembali informasi.

Cara mencari nilai *Recall* dalam kelas "Naik"

$$Recall = \frac{Jumlah\ data\ yang\ benar}{Jumlah\ data\ yang\ benar + Jumlah\ data\ yang\ salah}$$

$$Recall = \frac{838}{838 + 44} = 0,950$$

Untuk mencari nilai Cara mencari nilai *Recall* dalam kelas "Turun"

$$Recall = \frac{Jumlah\ data\ yang\ salah}{Jumlah\ data\ yang\ benar + Jumlah\ data\ yang\ salah}$$

$$Recall = \frac{172}{172 + 141} = 0,550$$

### 5. *F-Measure*

*F-Measure* merupakan gabungan perhitungan dari nilai *Precision* dan *Recall*.

Untuk mencari nilai *F-Measure* adalah:

$$F - Measure = \frac{2 * Precision * Recall}{Precision + Recall}$$

nilai *F-Measure* untuk data harga saham naik adalah:



$$F - Measure = \frac{2 * 0.856 * 0.950}{0.856 + 0.950} = 0.901$$

nilai *F-Measure* untuk data harga saham turun adalah:

$$F - Measure = \frac{2 * 0,796 * 0.550}{0,796 + 0.550} = 0.650$$

### Pengujian data *testing*

Data *testing* merupakan 20% dari jumlah data yang telah dikumpulkan. Jumlah data yang digunakan adalah sebanyak 1195, maka data yang digunakan sebagai data *testing* adalah sebanyak 239 yang diambil secara acak. Untuk melakukan pengujian kepada data *testing* yang belum diketahui kebenarannya, maka diperlukan beberapa proses sebagai berikut:

1. Siapkan data yang di *testing*, dan simpan dengan format file .arff,
2. Siapkan hasil data *training* yang akan dijadikan sebagai acuan dalam menentukan klasifikasi data *testing*,
3. Lakukan pengujian dengan menggunakan program WEKA.

### Pengujian Dengan Metode Naïve Bayes

Berikut adalah hasil dari pengujian data *testing* dengan program WEKA

```

Time taken to test model on training data: 0.5 seconds

=== Summary ===
Correctly Classified Instances      220          92.0502 %
Incorrectly Classified Instances    19           7.9498 %
Kappa statistic                    0.8222
Mean absolute error                 0.1206
Root mean squared error             0.2474
Relative absolute error             26.4415 %
Root relative squared error         51.8286 %
Total Number of Instances          239

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
          0.961   0.155   0.920     0.961   0.940     0.824  0.972   0.984   Naik
          0.845   0.039   0.922     0.845   0.882     0.824  0.972   0.958   Turun
Weighted Avg.   0.921   0.114   0.921     0.921   0.920     0.824  0.972   0.975

=== Confusion Matrix ===
 a  b  <-- classified as
149  6 | a = Naik
 13 71 | b = Turun
    
```

**Gambar 8** Hasil Prediksi WEKA Naïve Bayes

**Tabel 2** Klasifikasi Data *Testing* Naïve Bayes

<i>Prediction output</i>	Frekuensi	Total Jumlah Data
Benar	220	239
Salah	19	239

```

=== Predictions on test set ===

  inst#   actual  predicted error prediction
    1     1:?    1:Naik   0.98      0.98
    2     1:?    2:Turun   1         1
    3     1:?    1:Naik   0.989     0.989
    4     1:?    1:Naik   0.981     0.981
    5     1:?    1:Naik   0.985     0.985
    6     1:?    1:Naik   0.942     0.942
    7     1:?    2:Turun   0.812     0.812
    8     1:?    2:Turun   1         1
    9     1:?    1:Naik   0.733     0.733
   10     1:?    1:Naik   0.543     0.543
   11     1:?    1:Naik   0.975     0.975
   12     1:?    1:Naik   0.977     0.977
   13     1:?    1:Naik   0.984     0.984
   14     1:?    1:Naik   0.636     0.636
   15     1:?    1:Naik   0.634     0.634
   16     1:?    1:Naik   0.906     0.906
    
```

**Gambar 9** Contoh Hasil prediksi harga saham data *Testing* Naïve Bayes

Keterangan Gambar Gambar 9 adalah:

1. *Inst#* merupakan penomoran untuk setiap data yang ditest, *Actual* adalah output yang belum di isi hasilnya, dan hanya diisi dengan tanda tanya (?), *Predicted* merupakan klasifikasi output yang dihasilkan oleh program WEKA itu sendiri,
2. *Error prediction* merupakan presentasi jumlah kesalahan yang terjadi.

### Pengujian Dengan Metode *Decision Tree J-48*

Berikut adalah hasil dari pengujian data *training* dengan program WEKA.

```

=== Summary ===
Correctly Classified Instances   1182      98.9121 %
Incorrectly Classified Instances    13      1.0879 %
Kappa statistic                   0.9717
Mean absolute error                0.0205
Root mean squared error            0.1013
Relative absolute error            5.3074 %
Root relative squared error       23.0434 %
Total Number of Instances        1195

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
      -----  -
      0.997    0.032    0.989     0.997    0.993     0.972  0.989    0.992    Naik
      0.968    0.003    0.990     0.968    0.979     0.972  0.989    0.981    Turun
Weighted Avg.   0.989    0.024    0.989     0.989    0.989     0.972  0.989    0.989

=== Confusion Matrix ===

  a  b  <-- classified as
879  3  |  a = Naik
 10 303 |  b = Turun
    
```

**Gambar 10** Pengujian data *training* J-48

Berikut adalah hasil dari pengujian data *testing* dengan program WEKA

```

time taken to test model on training data: 0.04 seconds

=== Summary ===
Correctly Classified Instances      236          98.7448 %
Incorrectly Classified Instances     3           1.2552 %
Kappa statistic                     0.9724
Mean absolute error                 0.0238
Root mean squared error             0.1092
Relative absolute error             5.223 %
Root relative squared error        22.8629 %
Total Number of Instances          239

=== Detailed Accuracy By Class ===
                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
0.994   0.024   0.987   0.994   0.990   0.972   0.989   0.988   Naik
0.976   0.006   0.988   0.976   0.982   0.972   0.989   0.984   Turun
Weighted Avg.   0.987   0.018   0.987   0.987   0.987   0.972   0.989   0.986

=== Confusion Matrix ===
  a  b  <-- classified as
154  1  |  a = Naik
  2  82 |  b = Turun
    
```

**Gambar 11** Hasil Klasifikasi dengan Metode *Decision Tree* J-48

**Tabel 3** Klasifikasi Data *Testing*

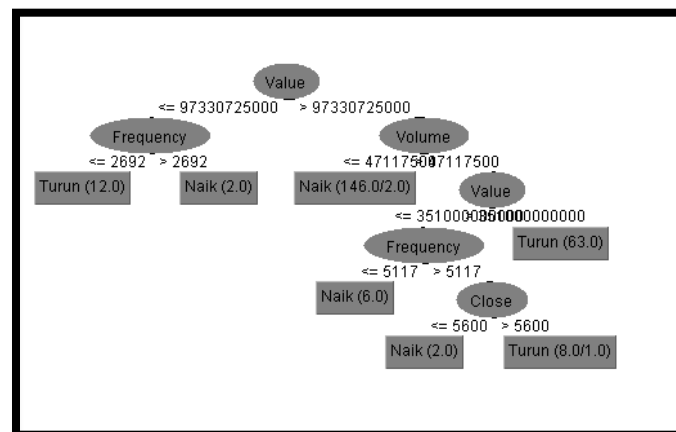
<i>Prediction output</i>	Frekuensi	Total Jumlah Data
Benar	236	239
Salah	3	239

```

=== Predictions on test set ===

inst#   actual   predicted  error  prediction
1       1:??    1:Naik    0.986
2       1:??    2:Turun    1
3       1:??    1:Naik    0.986
4       1:??    1:Naik    0.986
5       1:??    1:Naik    0.986
6       1:??    1:Naik    0.986
7       1:??    2:Turun    1
8       1:??    2:Turun    1
9       1:??    1:Naik    0.986
10      1:??    2:Turun    1
11      1:??    1:Naik    0.986
12      1:??    1:Naik    0.986
13      1:??    1:Naik    0.986
14      1:??    1:Naik    0.986
15      1:??    2:Turun    1
16      1:??    2:Turun    1
    
```

**Gambar 12** Contoh Hasil Prediksi *Decision Tree* J-48



**Gambar 13** Hasil Klasifikasi *Decision Tree* J-48

## Hasil Penelitian

**Tabel 4** Hasil Akurasi dan *Error Rate* pada Naïve Bayes

<i>Output Prediction</i>	Akurasi	<i>Error Rate</i>
<i>Data Training</i>	84.52%	15.48%
<i>Data Testing</i>	92.05%	7.95%

**Tabel 5** Hasil Akurasi dan *Error Rate* pada *Decision Tree* J-48

<i>Output Prediction</i>	Akurasi	<i>Error Rate</i>
<i>Data Training</i>	98.91%	1.09%
<i>Data Testing</i>	98.74%	1.26%

## Kesimpulan dan Saran

### Kesimpulan

Berdasarkan penelitian yang dilakukan penulis terhadap harga saham PT Astra Internasional Tbk, maka didapatkan kesimpulan bahwa:

1. Dengan menggunakan metode Naïve Bayes didapat hasil presentasi akurasi untuk data training sebesar 84.5188% dan untuk hasil data *testing* sebesar 92.0502 %.
2. Dengan keterangan 149 data harga saham yang diklasifikasikan mengalami kenaikan harga, dan terdapat 6 data harga saham yang seharusnya mengalami kenaikan harga saham, namun diklasifikasikan mengalami penurunan harga saham. Dan terdapat 71 data harga saham yang diklasifikasikan benar mengalami penurunan harga saham, namun terdapat 13 data harga saham yang seharusnya mengalami penurunan harga saham, tetapi diklasifikasikan mengalami kenaikan harga saham.
3. Dengan presentasi untuk nilai *precision* atau tingkat ketepatan informasi yang diharapkan oleh penulis dengan jawaban yang diberikan oleh sistem sebesar 0.920 dan nilai untuk *recall* atau tingkat keberhasilan terhadap informasi yang ditemukan kembali oleh sistem sebesar 0,961.
4. Sedangkan hasil prediksi dengan menggunakan metode *Decision Tree* J-48 didapatkan hasil presentasi akurasi untuk data training sebesar 98.9121 %, sedangkan hasil untuk akurasi pada data testing sebesar 98.7448 %.
5. Dengan keterangan 154 data harga saham yang diklasifikasikan mengalami kenaikan harga, dan terdapat 1 data harga saham yang seharusnya mengalami kenaikan harga, namun diklasifikasikan mengalami penurunan harga saham. Dan terdapat 82 data harga saham yang diklasifikasikan benar mengalami penurunan harga saham, namun terdapat 2 data harga saham yang seharusnya mengalami penurunan harga saham, tetapi diklasifikasikan mengalami kenaikan harga saham.
6. Dengan presentasi untuk nilai *precision* atau tingkat ketepatan informasi yang diharapkan oleh penulis dengan jawaban yang diberikan oleh sistem. sebesar 0.989 dan nilai untuk *recall* atau tingkat keberhasilan terhadap informasi yang ditemukan kembali oleh sistem sebesar 0,997.
7. Maka, harga saham yang diprediksi dengan menggunakan program WEKA untuk tanggal 4 Desember 2017 mengalami kenaikan harga.
8. Hasil ini juga didukung dengan kenaikan harga saham yang ada di *website* Bursa Efek Indonesia <http://ihsg-idx.com/saham/>, dengan kenaikan harga saham sebesar +125 atau +1.57%.

### Saran

Adapun saran bagi penelitian ini adalah:

1. Adanya pembuatan Website sehingga dapat melakukan prediksi harga saham dengan lebih cepat.
2. Menggunakan Metode pengerjaan yang berbeda, seperti penggunaan metode Jaringan Saraf Turuan, K-star, KNN.
3. Penambahan data untuk meningkatkan tingkat keberhasilannya.

### Referensi

1. Supranto. J, "Pasar Modal Indonesia. Rineka Cipta. Jakarta," Pasar Modal Indones., 1992.
2. D. S. Riswantoro, "Abstrak Penerapan Data Mining Untuk Memprediksikan Fluktuasi Harga Saham Menggunakan Metode Classification Dengan Teknik Dt ( Decision Tree )," pp. 4–5, 2011.
3. M. Ridwan, H. Suyono, and M. Sarosa, "Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *Eeccis*, vol. 7, no. 1, pp. 59–64, 2013.  
I. S. Achmad Basuki, "Decision Tree," p. 2.
4. M. S. H. Hasbullah, H Nasution, "No Title," 2015.
5. S. Fitri, "Perbandingan Kinerja Algoritma Klasifikasi Naïve Bayesian , Lazy-Ibk , Zero-R , Dan Decision Tree- J48," *Dasi*, vol. 15, no. 1, pp. 33–37, 2014.
6. D. Purnamasari, J. Henharta, Y. P. Sasmita, F. Ihsani, and I. W. S. Wicaksana, "Get Easy Using WEKA," 2013.