

ANALISIS KINERJA ALGORITMA J48 *DECISION TREE* UNTUK PENGAMBILAN KEPUTUSAN BELI/JUAL PADA SAHAM PT HARUM ENERGI Tbk. (HRUM)

(*PERFORMANCE ANALYSIS OF J48 DECISION TREE ALGORITHM FOR TAKING DECISION TO BUY / SELL IN STOCK PT HARUM ENERGI Tbk. (HRUM)*)

^[1]Nur Mishbah Hayat*, ^[2]Agung Budi Prasetyo, ^[3]Risma Septiana

^{[1][2][3]}Departemen Teknik Komputer
Universitas Diponegoro, Semarang, Indonesia

nmhayat@student.ce.undip.ac.id, agungprasetyo@ce.undip.ac.id, rismaseptiana@live.undip.ac.id

KEYWORDS:

Stocks, Technical Analysis, Investor, Data Mining, J48 Decision Tree, WEKA.

ABSTRACT

The problem that is often faced by investors in selling / buying stocks is the difficulty in analyzing a dataset of stock prices in large quantities. This analysis aims to predict the rise or fall of stock prices based on data obtained. To assist investors in determining buying / selling decisions on stock analysis based on technical and equipped with classification techniques in data mining. This study analyzes the performance of the J48 Decision Tree algorithm in the Waikato Environment Software for Knowledge Analysis (WEKA) version 3.8.2 for PT. Harum Energi Tbk. (HRUM). The J48 Decision Tree algorithm was chosen based on the results of previous studies which showed an accuracy percentage above 75%. The results showed in the testing data, the percentage of testing on data without normalization was higher by 87.3 (non-aggressive) and 88.8 (aggressive) compared to normalized data 84.2 (non-aggressive) and 85 (aggressive). The biggest stock profit generated is in non-aggressive type data without normalized by 48.75 or Rp. 48,750.00.

KATA KUNCI:

Saham, Analisis Teknikal, Investor, Data Mining, J48 Decision Tree, WEKA.

ABSTRAK

Masalah yang sering dihadapi oleh para investor dalam melakukan jual/beli saham adalah kesulitan dalam menganalisis suatu dataset harga saham yang jumlahnya sangat banyak. Analisis ini bertujuan untuk memprediksi naik atau turunnya harga saham berdasarkan data yang diperoleh. Untuk membantu investor dalam menentukan keputusan beli/jual pada saham berdasarkan analisis teknikal dan dikombinasikan dengan teknik klasifikasi pada data mining. Penelitian ini menganalisis kinerja dari algoritma J48 Decision Tree pada perangkat lunak Waikato Environment for Knowledge Analysis (WEKA) versi 3.8.2 untuk data harga saham PT. Harum Energi Tbk. (HRUM). Algoritma J48 Decision Tree dipilih berdasarkan hasil penelitian terdahulu yang menunjukkan persentase akurasi diatas 75%. Hasil penelitian menunjukkan bahwa pada pengujian data test, persentase akurasi pada data tanpa dinormalisasi lebih tinggi sebesar 87,3 (non agresif) dan 88,8 (agresif) dibandingkan dengan data yang dinormalisasi 84,2(non agresif) dan 85% (agresif). Keuntungan saham yang dihasilkan paling besar yaitu pada data tipe non agresif tanpa dinormalisasi sebesar 48,75 atau Rp 48.750,00.

I. PENDAHULUAN

Sebelum melakukan jual beli saham, *investor* atau *trader* perlu melakukan analisis terlebih dahulu. Analisis bertujuan untuk memprediksi naik atau turunnya harga saham berdasarkan data yang diperoleh. Analisis ini dibagi menjadi 2 yaitu,

analisis fundamental dan analisis teknikal. Analisis fundamental bertujuan untuk mengetahui nilai intrinsik saham. Sedangkan analisis teknikal merupakan teknik analisis yang digunakan untuk memprediksi tren suatu harga saham dengan cara mempelajari data pasar yang lampau, terutama

pergerakan harga dan volume. Analisis teknis dapat menggunakan berbagai model dan dasar, misalnya untuk pergerakan harga digunakan metode seperti indeks kekuatan relatif, indeks pergerakan rata-rata, regresi, korelasi antar pasar dan intra pasar, siklus, ataupun dengan cara klasik yaitu menganalisis pola grafik[1]. Analisis teknikal digunakan apabila investor atau trader ingin berinvestasi jangka pendek, dengan memperhatikan harga saham dan waktu (tren naik atau turun). Analisis teknikal yang digunakan meliputi indikator *Moving Average Convergence Divergence* (MACD), *Commodity Channel Index* (CCI), *Stochastic*, *Bollinger Bands*, *Pivot Point*, dan *Average Directional Index* (ADX).

Teknik klasifikasi merupakan salah satu teknik pada *Data Mining* yang berfokus pada pengelompokan sejumlah data kedalam kelas-kelas tertentu yang telah ditetapkan sebelumnya berdasarkan kesamaan sifat dan pola yang ada dalam data[2]. Salah satu algoritma klasifikasi adalah J48 *Decision Tree*. Algoritma J48 *Decision Tree* adalah sebuah algoritma turunan dari C4.5. Dasar dari algoritma ini adalah untuk membagi data ke dalam beberapa bagian berdasarkan nilai atribut dari item yang ada pada *training dataset*. Algoritma J48 *Decision Tree* dapat melakukan klasifikasi baik melalui *decision tree* ataupun *rules* yang diperoleh dari pohon tersebut[3]. Metodologi penelitian yang digunakan pada penelitian ini adalah *Knowledge Discovery in Databases* (KDD) *Process* yang terdiri dari 5 tahapan[4]. *Tools* yang digunakan untuk melakukan klasifikasi adalah *Waikato Environment for Knowledge Analysis* (WEKA). WEKA merupakan *software* terintegrasi yang berisi metode-metode *Data Mining*. WEKA dikembangkan oleh Universitas Waikato, Selandia Baru menggunakan bahasa pemrograman *Java*.

Salah satu pilihan saham adalah PT. Harum Energi Tbk dengan kode saham HRUM. Perusahaan ini terlibat dalam industri pertambangan, industri perdagangan, dan jasa. Saat ini, kegiatan bisnis utama perusahaan adalah beroperasi dan berinvestasi dalam penambangan dan logistik batubara melalui anak perusahaan[5].

Penelitian sebelumnya yang berkaitan dengan algoritma J48 *Decision Tree* antara lain adalah penelitian yang dilakukan oleh Tri Astuti dkk, (2016). Penelitian ini dilakukan guna memprediksi penyakit demam berdarah menggunakan metode algoritma J48 *Decision Tree* yang di

implementasikan pada data pasien menghasilkan nilai akurasi diatas 75%[3]. Pada contoh lainnya adalah penelitian yang dilakukan oleh Fergie Joanda Kaunang (2018) yang bertujuan untuk menganalisis tingkat kemiskinan menggunakan dimensi dasar dari Indeks Pembangunan Manusia (IPM) algoritma J48 *Decision Tree*. Menghasilkan nilai akurasi 88.6 % [6]. Penelitian selanjutnya dilakukan oleh Hendra Marcos dan Indriana Hidayah (2014) terkait implementasi *Data Mining* untuk klasifikasi nasabah kredit bank menggunakan *Classification Rule*. Tujuan penelitian ini adalah memberikan kontribusi untuk mempercepat pengambilan keputusan tersebut. Teknik *Data Mining* digunakan untuk memberikan model agar pihak bank dapat lebih cepat dalam membuat keputusan kepada nasabah yang berhak diberi kredit atau ditolak. Disimpulkan dari beberapa algoritma bahwa algoritma C4.5 yang mempunyai nilai akurasi yang tertinggi[7]. Penelitian lainnya adalah penelitian yang dilakukan oleh Nurfaizah, dkk (2017). Penelitian ini berkaitan dengan pemanfaatan algoritma *Decision Tree*-J48, *K-Nearest*, dan *Zero-R* pada kinerja akademik. Tujuan pada penelitian ini adalah mendeteksi dengan teknik *Data Mining* agar dapat mencatat berbagai masalah pada siswa. Penelitian ini membandingkan tingkat akurasi 3 algoritma tersebut untuk evaluasi kinerja akademik siswa. Dari dataset yang sama, penggunaan *Decision Tree*-J48, *K-Nearest*, dan *Zero-R* sama-sama mempunyai nilai akurasi diatas 85% [8].

Penelitian selanjutnya yang berkaitan dengan indikator pada analisis teknikal antara lain yang dilakukan oleh M. Lutfi Sulthon (2018). Penelitian ini dilakukan guna mendeteksi sinyal jual/beli saham dengan menggunakan *fuzzy rule-based evidential reasoning* dan *c-means clustering*. *Fuzzy rule-based evidential reasoning* dapat memetakan kondisi pergerakan saham sedangkan *clustering* digunakan untuk menentukan jumlah fungsi keanggotaan masing-masing atribut. Dari pengujian sistem diperoleh ROI sebesar 83,80% dari profit maksimal yang didapatkan selama pengujian[1]. Penelitian lainnya dilakukan oleh Mutmainah dan Sri Sulasmiyati (2017) yang bertujuan untuk menguji keakuratan analisis teknikal indikator *Stochastic Oscillator* dalam menentukan sinyal beli dan jual pada saham yang tergabung dalam sub sektor konstruksi dan bangunan periode 2014-2016. Hasil penelitian diperoleh bahwa indikator *Stochastic*

Oscillator dikatakan akurat dengan menghasilkan 269 sinyal beli dan jual dalam kurun waktu 2014-2017[9].

Berdasarkan penelitian-penelitian sebelumnya, dapat disimpulkan bahwa penggunaan Algoritma J48 *Decision Tree* untuk analisis saham pada PT Harum Energi Tbk. akan menjadi solusi yang efektif.

II. METODE PENELITIAN

Metodologi penelitian yang digunakan pada penelitian ini adalah Knowledge Discovery in Databases (KDD) Process. Metode ini terdiri dari 5 tahap yaitu data selection, data preprocessing, data transformation, data mining, dan interpretation/evaluation.

A. Data Selection

Data penelitian yang akan diteliti adalah data harga saham PT. Harum Energi (HRUM) Tbk. Data harga saham yang digunakan memiliki rentang waktu selama 2,5 tahun mulai dari tanggal 1 Juni 2016 sampai dengan tanggal 31 Desember 2018 didapatkan dengan cara mengunduh dari situs resmi *Yahoo Finance* dengan mengakses laman <https://finance.yahoo.com>. Data tersebut berupa dokumen .csv yang berisi *timestamp* atau keterangan waktu, nilai *Open*, *High*, *Low*, *Close*, *adjusted close*, dan *volume* transaksi saham yang selalu diperbarui setiap harinya. Untuk mengolah data tersebut, perlu dilakukan perubahan ekstensi dokumen dari .csv menjadi .xlsx.

B. Data Preprocessing

Data harga saham yang terdiri dari nilai *Open*, *High*, *Low*, *Close*, *adjusted close* dan *volume* kemudian diolah dengan menambahkan 6 perhitungan indikator, yaitu *Moving Average Convergence Divergence* (MACD), *Commodity Channel Index* (CCI), *Stochastic*, *Bollinger Band*, *Pivot Point*, dan *Average Directional Index* (ADX)[10]. Perhitungan ini dimaksudkan untuk mendapatkan informasi yang membantu proses klasifikasi pada data mining.

Data kemudian dibagi menjadi 2 bagian, yakni data training dan data test. Data training digunakan untuk melatih model dengan algoritma, sedangkan data test digunakan untuk mengetahui akurasi prediksi dari model yang sudah didapatkan dari data training.

Dalam penelitian ini terdapat dua macam data yang digunakan untuk pembandingan pada tahap analisis, antara lain:

1. Berdasarkan *Trading Style*

Data training berdasarkan *trading style* dibedakan menjadi dua, yaitu non agresif dan agresif. Keputusan beli pada *trading style* non agresif adalah saat grafik saham melewati garis batas bawah pada indikator *Bollinger Band*. Sedangkan keputusan jual dilakukan pada saat grafik saham melewati garis batas atas pada indikator *Bollinger Band*. Pada *trading style* non agresif terdapat 16 titik beli/jual dapat dilihat pada Gambar 1.



Gbr. 1 Penentuan sinyal non agresif

Sedangkan pada *trading style* agresif digunakan 3 indikator, yaitu *Bollinger Band*, *MACD* dan *Stochastic*. Pada indikator *Bollinger Band*, keputusan beli dilakukan ketika grafik harga saham mendekati batas bawah dan keputusan jual dilakukan ketika grafik harga saham mendekati batas atas. Pada indikator *MACD* keputusan beli dilakukan jika garis *MACD* bersinggungan dengan garis sinyal kearah atas (*golden cross*). Sedangkan keputusan jual dilakukan jika garis *MACD* bersinggungan dengan garis sinyal kearah bawah (*death cross*). Pada indikator *Stochastic* keputusan beli dilakukan saat grafik melewati atau berada pada angka 80 sedangkan keputusan beli dilakukan saat grafik melewati atau berada pada angka 20. Pada *trading style* agresif terdapat 24 titik beli/jual yang ditunjukkan pada Gambar 2.



Gbr. 2 Penentuan sinyal agresif

2. Normalisasi data

Data training berdasarkan proses *preprocessing* dibedakan menjadi dua, yaitu normalisasi dan tanpa

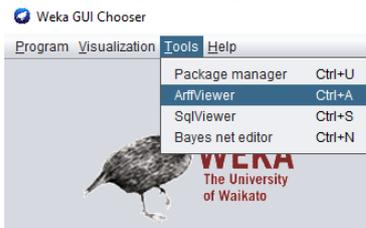
normalisasi. Data yang di normalisasi merupakan perubahan rentang nilai data hasil perhitungan dari data tanpa normalisasi. Nilai yang dihasilkan minimal 0 dan maksimal 1. Proses normalisasi dapat dilakukan dengan menggunakan fitur normalisasi pada perangkat lunak WEKA. Sedangkan pada data tanpa normalisasi merupakan data asli dari hasil perhitungan yang sudah dilakukan sebelumnya. Perbedaan data sebelum dan sesudah normalisasi ditampilkan pada Gbr. 3.

	A	B	C	D	E	F	G		A	B	C	D	E	F	G	
1	UB	MB	LB	PP	R1	R2	R3	1	UB	MB	LB	PP	R1	R2	R3	
2	0	0	0	0	0	0	0	2	0	0	0	0	818.33	826.67	833.33	841.67
3	0	0	0	0.010059	0.022204	0.028662	0.038202	3	0	0	0	0	838.33	871.67	893.33	926.67
4	0	0	0	0.03437	0.047894	0.054143	0.063667	4	0	0	0	0	886.67	923.33	946.67	983.33
5	0	0	0	0.052808	0.059211	0.062102	0.065169	5	0	0	0	0	923.33	946.67	963.33	986.67
6	0	0	0	0.059517	0.062497	0.063697	0.063967	6	0	0	0	0	936.67	953.33	966.67	983.33
7	0	0	0	0.055162	0.055837	0.067877	0.066663	7	0	0	0	0	930	945	975	990
8	0	0	0	0.057002	0.06003	0.063697	0.063667	8	0	0	0	0	931.67	948.33	966.67	983.33
9	0	0	0	0.078794	0.095393	0.108282	0.118351	9	0	0	0	0	975	1020	1060	1105
10	0	0	0	0.160103	0.225326	0.257168	0.304117	10	0	0	0	0	1196.67	1283.33	1371.67	1518.33
11	0	0	0	0.210397	0.240129	0.297775	0.310858	11	0	0	0	0	1296.67	1313.33	1404.67	1533.33
12	0	0	0	0.186027	0.186094	0.187995	0.181047	12	0	0	0	0	1150	1165	1185	1200
13	0	0	0	0.202853	0.232728	0.247614	0.263667	13	0	0	0	0	1221.67	1298.33	1351.67	1428.33
14	0	0	0	0.202013	0.208881	0.227708	0.223969	14	0	0	0	0	1220	1250	1310	1340

Gbr. 3 Proses normalisasi

C. Data Transformation

Tranformasi atau perubahan bentuk pada data agar sesuai untuk dilanjutkan ke proses *data mining*. Transformasi data meliputi perubahan bentuk yang semula data saham berupa *string* ke dalam bentuk numerik atau nominal dan perubahan format dokumen dari *.csv* diubah menjadi *.arff* agar kompatibel dengan perangkat lunak WEKA dengan menggunakan fitur pada perangkat lunak WEKA yaitu *ArffViewer* ditunjukkan pada Gambar 4.



Gbr. 4 Fitur *ArffViewer* pada WEKA

D. Data Mining

Tahap ini berupa proses mencari pola dalam data terpilih dengan menggunakan teknik atau metode tertentu. Data yang sudah dilakukan transformasi kemudian akan diolah dengan perangkat lunak WEKA untuk dicari polanya menggunakan algoritma (*classifier*) *data mining* yang digunakan. Proses pelatihan dan pengujian data dalam perangkat lunak WEKA terdiri dari dua tahap, yaitu:

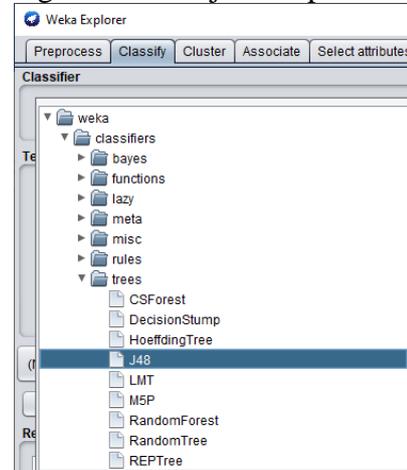
1. Prapengolahan

Tahap prapengolahan berguna untuk menghapus atribut pada data yang tidak dibutuhkan. Bertujuan agar atribut yang tidak dibutuhkan tidak mempengaruhi proses maupun hasil klasifikasi. 21 atribut yang dibutuhkan adalah BB atas, BB tengah,

BB Bawah, PP, R1, R2, R3, S1, S2, S3, MACD, Signal, Fast %K, Slow %D, Smoothed %D, +DI14, -DI14, ADX, CCI, MA, CLASS.

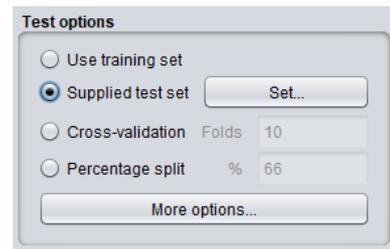
2. Klasifikasi

Tahap klasifikasi merupakan pemilihan algoritma (*classifier*) untuk mencari pola pada data *training*. Pemilihan algoritma ditunjukkan pada Gambar 5.

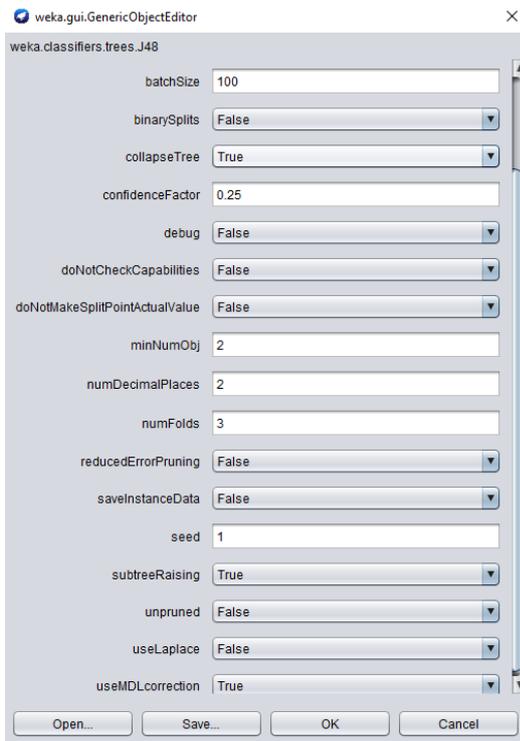


Gbr. 5 Memilih Algoritma J48

Pada proses *data mining* dihasilkan model algoritma J48 *Decision Tree* dari *data training* menggunakan *test options* dengan memilih *Use training set*. Sedangkan menguji *data testing* menggunakan *test options* dengan memilih *Supplied test set*. Tampilan *test option* ditunjukkan pada Gambar 6.



Gbr. 6 *Test options* pada WEKA



Gbr. 7 Parameter pada classifier J48

Algoritma atau *classifier* J48 memiliki beberapa parameter seperti yang ditampilkan pada Gambar 7. Pada penelitian ini digunakan parameter dengan nilai awal yang sudah ditetapkan oleh WEKA.

Pada algoritma J48, untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan persamaan 1 seperti berikut[11]:

$$Gain(S, A) = S + \sum_{i=1}^n \frac{|S_i|}{|S|} * S_i \quad (1)$$

Sementara itu, penghitungan nilai entropi dapat dilihat pada persamaan 2 seperti berikut[11]:

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \quad (2)$$

Keterangan :

- S : himpunan kasus
- A : fitur
- n : jumlah partisi S
- |S_i| : jumlah kasus pada partisi ke-i
- |S| : jumlah kasus dalam S
- p_i : proporsi dari S_i terhadap S

E. Interpretation/Evaluation

Hasil dari pelatihan data *training* menghasilkan persentase akurasi seberapa tepat kelas yang sudah diberikan pada data *training*. Gambar 11 merupakan tampilan dari hasil klasifikasi data latih menggunakan *classifier* J48.

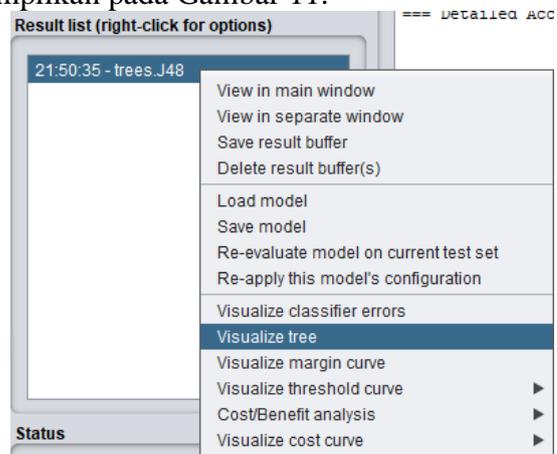
Pada data *test* juga dihasilkan klasifikasi. Klasifikasi dihasilkan oleh model yang sudah dilatih

dengan data *training*. Juga terdapat *actual class* yang sudah diberikan agar dapat melihat perbedaan sinyal (manual) dengan klasifikasi yang dihasilkan oleh WEKA. Hasil klasifikasi ditampilkan pada Gambar 8.

Classifier output				
=== Predictions on test set ===				
inst#	actual	predicted	error	prediction
1	1:HOLD	1:HOLD	0.985	
2	1:HOLD	1:HOLD	0.985	
3	1:HOLD	1:HOLD	0.985	
4	1:HOLD	1:HOLD	0.985	
5	2:SELL	2:SELL	1	
6	2:SELL	2:SELL	1	
7	2:SELL	2:SELL	1	
8	2:SELL	2:SELL	1	
9	2:SELL	1:HOLD	+	0.985
10	1:HOLD	1:HOLD	0.985	
11	1:HOLD	1:HOLD	0.985	
12	2:SELL	1:HOLD	+	0.985
13	1:HOLD	1:HOLD	0.985	
14	1:HOLD	1:HOLD	0.985	
15	2:SELL	1:HOLD	+	0.985
16	2:SELL	1:HOLD	+	0.985

Gbr. 8 Hasil klasifikasi data test

Pohon keputusan dihasilkan dengan klik kanan pada model yang didapatkan, selanjutnya pilih *visualize tree*. Langkah ini ditampilkan pada Gambar 10 dan pohon keputusan yang dihasilkan ditampilkan pada Gambar 11.



Gbr. 9 Langkah visualize tree

1. Pengukuran Kinerja

Kinerja dari algoritma J48 *Decision Tree* untuk pengambilan keputusan beli/jual pada saham PT. Harum Energi Tbk (HRUM) dapat dilihat dari jumlah keuntungan dan persentase akurasi. Persentase akurasi terdapat pada *summary* setelah melakukan *training* maupun *test*. Persentase akurasi tersebut selanjutnya akan dibandingkan antara satu dengan yang lain sehingga diketahui model mana yang memiliki kinerja terbaik.

Perhitungan keuntungan dilakukan untuk mengetahui keuntungan yang didapatkan dari proses beli/jual saham selama 1 tahun berdasarkan sinyal hasil klasifikasi. Terdapat tiga rumus dasar untuk

menghitung keuntungan dan kinerja dari algoritma atau classifier antara lain:

a. Kondisi sinyal beli

Ketika sinyal beli, maka diasumsikan untuk membeli saham dengan modal awal sebanyak Rp. 100.000,- . Perhitungan pada kondisi ini dapat menggunakan persamaan 1.

$$Beli = \frac{Modal\ Awal}{Close} \quad (1)$$

b. Kondisi sinyal jual

Ketika sinyal jual, maka diasumsikan untuk menjual saham yang dimiliki sesuai dengan harga close pada hari tersebut. Perhitungan pada kondisi ini dapat menggunakan persamaan 2.

$$Jual = Asset \times Close \quad (2)$$

c. Persentase Keuntungan

Perhitungan presentase keuntungan dilakukan agar dapat mengetahui berapa persen keuntungan trading saham selama 1 tahun. Perhitungan persentase keuntungan dapat menggunakan persamaan 3.

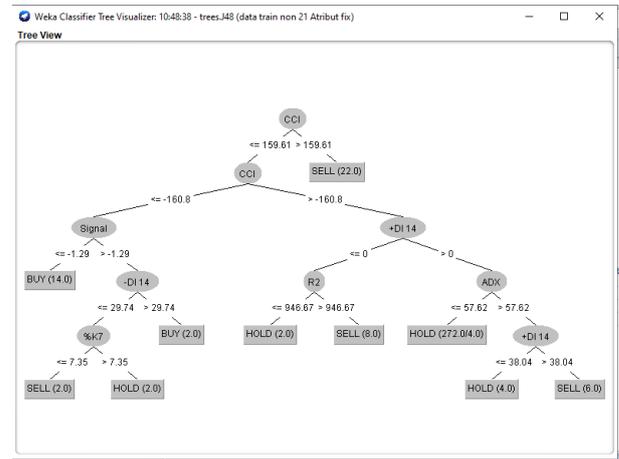
$$\% \text{Keuntungan} = \frac{Uang - Modal\ Awal}{Modal\ Awal} \times 100 \% \quad (3)$$

III. HASIL PENELITIAN DAN PEMBAHASAN

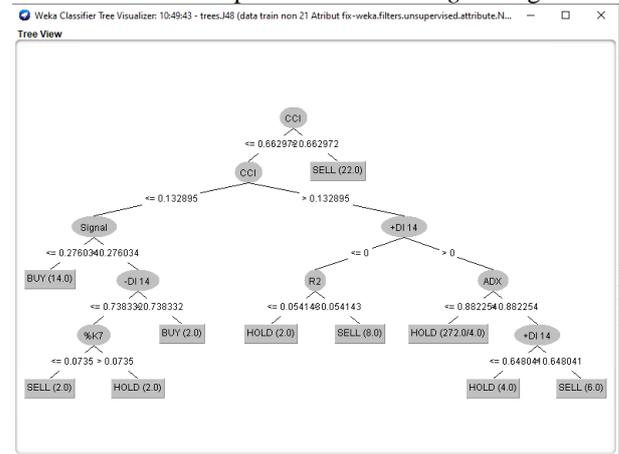
1. Pelatihan Model dan Hasil Klasifikasi pada Data Training

F. Data Non Agresif

Pelatihan model dengan data training non agresif dilakukan pada dua jenis data yaitu data normalisasi dan tanpa normalisasi. Hasil dari proses pelatihan kedua jenis data tersebut dengan menggunakan algoritma J48 Decision Tree tidak memperlihatkan perbedaan. Pohon keputusan maupun persentase akurasi yang dihasilkan memperlihatkan hasil yang sama, kecuali nilai pada pohon yang dinormalisasi hanya memiliki rentang nilai dari 0 sampai dengan 1.



Gbr. 10 Pohon keputusan data training non agresif



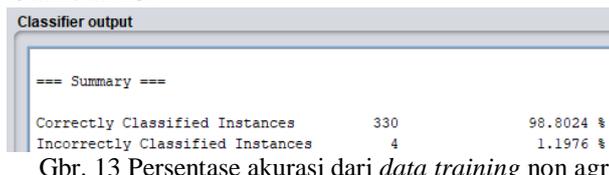
Gbr. 11 Pohon keputusan data training non agresif normalisasi

Gambar 10 dan 11 merupakan pohon keputusan yang dihasilkan dari pelatihan data training non agresif. Pemilihan atribut sebagai akar (root) maupun cabang didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Sedangkan angka yang menjadi nilai atribut pada cabang diperoleh dari data pada kelas (keputusan) tertentu. 159,61 merupakan angka batas terkecil pada atribut CCI yang kelas (keputusan) bernilai sell. Jika CCI bernilai kurang dari 159,61 penentuan keputusan dilihat atribut lain berdasarkan pada nilai gain tertinggi. Pembentukan pohon akan berhenti sampai terbentuk leaf dengan keputusan akhir (buy, sell, hold).

Sebanyak 7 atribut dengan Gain terbaik muncul pada pohon keputusan dari total 21 atribut. Atribut tersebut antara lain CCI, Signal, +DI 14, -DI 14, R2, ADX, dan %K7. Pohon keputusan yang dihasilkan pada data tanpa normalisasi dan normalisasi tidak ada perbedaan.

Pelatihan model pada perangkat lunak WEKA menunjukkan bahwa data yang diklasifikasikan sesuai dengan sinyal yang diberikan sebanyak 330 data atau 98,8%. Sedangkan data yang

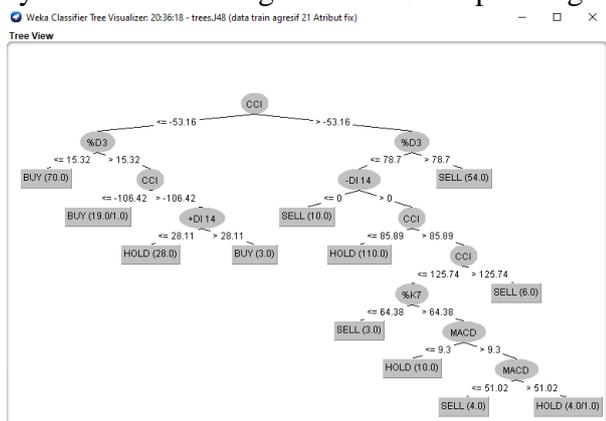
diklasifikasikan tidak sesuai sebanyak 4 data atau 1.2%. Hasil persentase tersebut diperlihatkan pada Gambar 13.



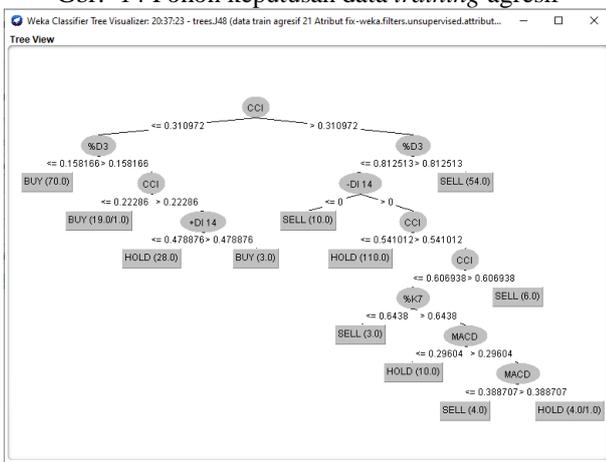
Gbr. 13 Persentase akurasi dari data training non agresif

G. Data Tipe Agresif

Pelatihan model dengan data training agresif dilakukan pada dua jenis data yaitu data normalisasi dan tanpa normalisasi. Hasil dari proses pelatihan kedua jenis data tersebut dengan menggunakan algoritma J48 Decision Tree tidak memperlihatkan perbedaan. Pohon keputusan maupun persentase akurasi yang dihasilkan memperlihatkan hasil yang sama, kecuali nilai pada pohon yang dinormalisasi hanya memiliki rentang nilai dari 0 sampai dengan 1.



Gbr. 14 Pohon keputusan data training agresif



Gbr. 15 Pohon keputusan data training agresif normalisasi

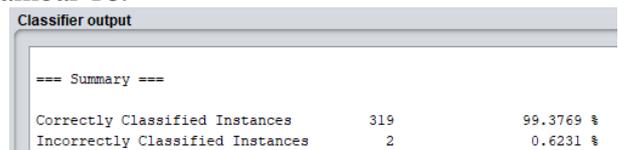
Gambar 14 dan 15 merupakan pohon keputusan yang dihasilkan dari pelatihan data training agresif. Pemilihan atribut sebagai simpul, baik akar (root) atau simpul internal didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Sedangkan angka yang menjadi nilai atribut pada cabang

diperoleh dari data pada kelas (keputusan) tertentu. -53,16 merupakan angka batas terkecil pada atribut CCI dengan keputusan bernilai antara hold dan sell. Jika CCI bernilai kurang dari sama dengan -53,16 maka keputusan akhirnya bernilai antara hold dan buy. Pembentukan pohon akan berhenti sampai terbentuk leaf yang berupa keputusan (buy, sell, hold).

Sebanyak 6 atribut dengan Gain tertinggi muncul pada pohon keputusan dari total 21 atribut. Atribut tersebut antara lain CCI, %D3, +DI 14, -DI 14, %K7 dan MACD. Pohon keputusan yang dihasilkan pada data tanpa normalisasi dan normalisasi tidak ada perbedaan.

Pelatihan model dengan data training agresif dilakukan pada dua jenis data yaitu data normalisasi dan tanpa normalisasi. Hasil dari proses pelatihan kedua jenis data tersebut dengan menggunakan algoritma J48 Decision Tree tidak memperlihatkan perbedaan. Persentase akurasi maupun pohon keputusan yang dihasilkan memperlihatkan hasil yang sama.

Pelatihan model pada perangkat lunak WEKA menunjukkan bahwa data yang diklasifikasikan sesuai dengan sinyal yang diberikan sebanyak 319 data atau 99,4%. Sedangkan data yang diklasifikasikan tidak sesuai sebanyak 2 data atau 0.6%. Hasil persentase tersebut diperlihatkan pada Gambar 16.



Gbr. 16 Persentase akurasi data training agresif

2. Pengujian Model dan Hasil Klasifikasi pada Data Test

A. Data Tipe Non Agresif

Pengujian model yang sudah diperoleh pada data test non agresif dilakukan pada dua jenis data yaitu data tanpa normalisasi dan normalisasi. Pohon keputusan yang dihasilkan sama dengan yang dihasilkan pada pelatihan data non agresif sesuai dengan Gambar 12. Pengujian jenis data tanpa normalisasi pada perangkat lunak WEKA menunjukkan bahwa data yang diklasifikasikan sesuai dengan sinyal yang diberikan sebanyak 228 data atau 87,4%. Sedangkan data yang diklasifikasikan tidak sesuai sebanyak 33 data 12.6%.

Hasil persentase tersebut diperlihatkan pada Gambar 17.

Classifier output			
=== Summary ===			
Correctly Classified Instances	228	87.3563 %	
Incorrectly Classified Instances	33	12.6437 %	

Gbr. 17 Persentase akurasi data *test* non agresif non normalisasi

Sedangkan pengujian jenis data normalisasi pada perangkat lunak *WEKA* menunjukkan bahwa data yang diklasifikasikan sesuai dengan sinyal yang diberikan sebanyak 84,3%. Sedangkan data yang diklasifikasikan tidak sesuai sebanyak 15,7%. Hasil persentase tersebut diperlihatkan pada Gambar 18.

Classifier output			
=== Summary ===			
Correctly Classified Instances	220	84.2912 %	
Incorrectly Classified Instances	41	15.7088 %	

Gbr. 18 Persentase akurasi data *test* non agresif normalisasi

Pada data non agresif diperoleh hasil persentase akurasi lebih tinggi pada data tanpa dinormalisasi sebesar 87,4% dibandingkan dengan data yang dinormalisasi sebesar 84,3%.

B. Data Tipe Agresif

Pengujian model yang sudah diperoleh pada data *test* agresif dilakukan pada dua jenis data yaitu data tanpa normalisasi dan normalisasi. Pohon keputusan yang dihasilkan sama dengan yang dihasilkan pada pelatihan data agresif sesuai dengan Gambar 13. Pengujian jenis data tanpa normalisasi pada perangkat lunak *WEKA* menunjukkan bahwa data yang diklasifikasikan sesuai dengan sinyal yang diberikan sebanyak 231 data atau 88,8%. Sedangkan data yang diklasifikasikan tidak sesuai sebanyak 29 data atau 11,2%. Hasil persentase tersebut diperlihatkan pada Gambar 19.

Classifier output			
=== Summary ===			
Correctly Classified Instances	231	88.8462 %	
Incorrectly Classified Instances	29	11.1538 %	

Gbr. 19 Persentase akurasi data *test* agresif tanpa normalisasi

Sedangkan pengujian jenis data normalisasi pada perangkat lunak *WEKA* menunjukkan bahwa data yang diklasifikasikan sesuai dengan sinyal yang diberikan sebanyak 221 data atau 85%. Sedangkan data yang diklasifikasikan tidak sesuai sebanyak 39 data atau 15%. Hasil persentase tersebut diperlihatkan pada Gambar 20.

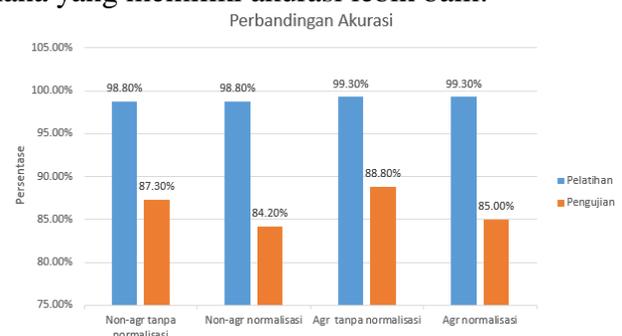
Classifier output			
=== Summary ===			
Correctly Classified Instances	221	85 %	
Incorrectly Classified Instances	39	15 %	

Gbr. 20 Persentase akurasi data *test* agresif normalisasi

Pada data agresif diperoleh hasil persentase akurasi lebih tinggi pada data tanpa dinormalisasi sebesar 88,8% dibandingkan dengan data yang dinormalisasi sebesar 85%.

3. Perbandingan Akurasi

Persentase yang telah didapatkan dari 2 pelatihan data *training* dan 2 pengujian data *test* dapat divisualisasikan dengan menggunakan grafik, berguna untuk memperlihatkan perbandingan antara data satu dengan yang lain dan mengetahui model mana yang memiliki akurasi lebih baik.



Gbr. 21 Grafik perbandingan akurasi setiap model

Gambar 21 menunjukkan perbandingan persentase akurasi pada setiap model. Terlihat bahwa data *training* tanpa normalisasi dan normalisasi tidak menunjukkan perbedaan pada nilai persentase akurasi, yaitu 98,8% pada tipe non agresif dan 99,3% pada tipe agresif. Sedangkan pada data *test* menunjukkan persentase yang berbeda pada setiap model dan mengalami penurunan, yaitu 87,3% pada tipe non agresif tanpa normalisasi dan 84,2% pada tipe non agresif normalisasi, 88,8% pada tipe agresif tanpa normalisasi dan 85% pada tipe agresif normalisasi.

Akurasi tertinggi data *training* diperoleh pada model dengan tipe data agresif baik tanpa normalisasi maupun normalisasi dengan 99,3%. Sedangkan akurasi tertinggi data *test* diperoleh pada model dengan tipe data agresif tanpa normalisasi dengan 88,8%.

4. Perbandingan Keuntungan

Hasil berupa klasifikasi sinyal yang diperoleh dari pengujian pada data *test* dapat digunakan sebagai acuan untuk menghitung keuntungan saham dalam rentang waktu satu tahun. Kemudian keuntungan

dari hasil klasifikasi akan dibandingkan dengan keuntungan yang didapatkan dari penentuan keputusan beli/jual secara manual.

Pada data uji tipe non agresif tanpa normalisasi memiliki keuntungan paling tinggi sebesar 48,75%. 48,75% diperoleh dari persamaan 4 yang mengacu pada persamaan 3.

$$\% \text{keuntungan} = \frac{148700 - 100.000}{100.000} \times 100 \% \\ = 48,75\% \quad (4)$$

Sedangkan data uji tipe data non agresif normalisasi mendapatkan keuntungan 8,92%. 8,92% diperoleh dari persamaan 5 yang mengacu pada persamaan 5.

$$\% \text{keuntungan} = \frac{108920 - 100.000}{100.000} \times 100 \% \\ = 8,92\% \quad (5)$$

Selanjutnya data uji agresif tanpa normalisasi memiliki keuntungan 26,20%. 26,20% diperoleh dari persamaan 4 yang mengacu pada persamaan 6.

$$\% \text{keuntungan} = \frac{126200 - 100.000}{100.000} \times 100 \% \\ = 26,20\% \quad (6)$$

Data uji tipe agresif normalisasi mendapatkan keuntungan sebesar 2,68% yang merupakan keuntungan paling kecil. 2,68% diperoleh dari persamaan 7 yang mengacu pada persamaan 3.

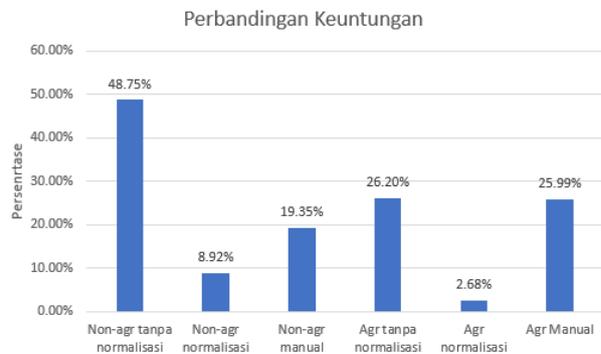
$$\% \text{keuntungan} = \frac{102680 - 100.000}{100.000} \times 100 \% \\ = 2,68\% \quad (7)$$

Selain menghitung keuntungan saham berdasarkan klasifikasi yang dihasilkan perangkat lunak WEKA, Penulis juga menghitung keuntungan saham dari kelas yang diberikan secara manual pada data *test*. Hasil keuntungan saham dari tipe non agresif sebesar 19,35%. 19,35% diperoleh dari persamaan 8 yang mengacu pada persamaan 3.

$$\% \text{keuntungan} = \frac{119340 - 100.000}{100.000} \times 100 \% \\ = 19,34\% \quad (8)$$

Hasil keuntungan saham dari tipe agresif sebesar 25,99%. 25,99% diperoleh dari persamaan 9 yang mengacu pada persamaan 3.

$$\% \text{keuntungan} = \frac{125990 - 100.000}{100.000} \times 100 \% \\ = 25,9\% \quad (9)$$



Gbr. 22 Grafik persentase keuntungan saham

Gambar 22 menunjukkan perbandingan persentase keuntungan pada setiap data uji. Terlihat bahwa data tanpa normalisasi mendapatkan keuntungan yang lebih tinggi dibanding dengan data yang dinormalisasi. Pada tipe data non agresif tanpa normalisasi mendapat keuntungan sebesar 48,75%, sedangkan pada data yang dinormalisasi mendapat keuntungan sebesar 8,92%. Selanjutnya keuntungan yang didapatkan pada tipe data agresif tanpa normalisasi sebesar 26,2%, sedangkan pada data yang dinormalisasi mendapat keuntungan sebesar 2,68%.

IV. KESIMPULAN

Berdasarkan penelitian yang dilakukan terhadap data harga saham PT. Harum Energi Tbk. (HRUM) menggunakan algoritma *J48 Decision Tree*, dapat disimpulkan bahwa hasil persentase akurasi data *test* tipe non agresif lebih tinggi pada data tanpa normalisasi sebesar 87,3% dibandingkan data yang di normalisasi yaitu sebesar 84,2%. Hasil persentase akurasi data *test* tipe agresif lebih tinggi pada data tanpa normalisasi sebesar 88,8% dibandingkan data yang di normalisasi yaitu sebesar 85%. Hasil persentase akurasi tertinggi data *test* dimiliki oleh data tipe agresif tanpa normalisasi sebesar 88,8%. Persentase keuntungan paling tinggi dimiliki oleh data tipe agresif tanpa normalisasi sebesar 48,75%. Persentase keuntungan pada data tanpa normalisasi lebih besar dibandingkan data yang di normalisasi. Persentase keuntungan pada kelas hasil klasifikasi (tanpa normalisasi) lebih tinggi dibandingkan pada kelas yang diberikan manual. Sebesar 48,75% pada tipe data non agresif dan 26,2% pada tipe data agresif. Sedangkan pada kelas manual sebesar 19,35% untuk tipe agresif dan 25,99% untuk tipe non agresif. Pada saham PT Harum Energi Tbk. (HRUM) lebih baik menggunakan *trading style* non agresif untuk mendapatkan keuntungan optimal dengan syarat data tidak di normalisasi.

UCAPAN TERIMA KASIH

Terima kasih disampaikan kepada dosen pembimbing dan semua pihak yang telah meluangkan waktu untuk membantu menyelesaikan penelitian ini

REFERENSI

- [1] M. L. S. A.S., A. B. Prasetijo, and M. Somantri, "Pendeteksi Sinyal Jual/Beli Saham dengan *Fuzzy Rule-Based Evidential Reasoning* dan *C-means Clustering*," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 7, no. 3, 2018.
- [2] J. Han, M. Kamber, and J. Pei, *Data mining: Data mining concepts and techniques, Third*. New York: Morgan Kaufmann, 2014.
- [3] T. Astuti, I. Mujiati, D. Ayu, V. Ristianah, and W. A. Lestari, "Penerapan Algoritme J48 Untuk Prediksi," *J. Telemat.*, vol. 9, no. 2, pp. 1–10, 2016
- [4] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From Data Mining to Knowledge Discovery in Databases," *AI Mag.*, vol. 17, no. 3, pp. 37–54, 1996.
- [5] PT Harum Energi, "Profil Manajemen". [Online] Available: <http://www.harumenergy.com/id/about/25/profil-manajemen> [Accessed: 3 Agustus 2019].
- [6] F. J. Kaunang, "Penerapan Algoritma J48 Decision Tree Untuk Analisis Tingkat Kemiskinan di Indonesia," *CogITo Smart J.*, vol. 4, no. 2, p. 348, 2019. M. Shell. (2002) IEEEtran homepage on CTAN. [Online]. Available: <http://www.ctan.org/tex-archive/macros/latex/contrib/supported/IEEEtran/>
- [7] H. Marcos and I. Hidayah, "Implementasi Data Mining Untuk Klasifikasi Nasabah Kredit Bank " X " Menggunakan Classification Rule," *Semin. Nas. Teknol. Inf. dan Multimed.*, pp. 1–7, 2014.
- [8] Nurfaizah, M. Imron, and L. Perdanawati, "Algoritma decision tree-j48, k-nearest, dan zero-r pada kinerja akademik," *Semin. Nas. Teknol. Inf.*, pp. 12–18, 2017.
- [9] Mutmainah and S. Sulasmiyati, "Analisis Teknikal Indikator *Stochastic Oscillator* dalam Menentukan Sinyal Beli Dan Sinyal Jual Saham," *J. Adm. Bisnis*, vol. 49, no. 1, pp. 1–8, 2017.
- [10] *Stockcharts*, "Technical Indicators and Overlays". [Online] Available: https://school.stockcharts.com/doku.php?id=technical_indicators [Accessed: 9 Agustus 2019].
- [11] Kusri and E. T. Lutfi, *Algoritma Data Mining*. Yogyakarta: Penerbit Andi, 2009.