

Prediksi Pergerakan Volatilitas Pasar Saham Berdasarkan *Financial News*

Nadhieya Safiera ZD¹, Deni Saepudin², Aniq Atiqi Rohmawati³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹nadhieyasafiera@students.telkomuniversity.ac.id, ²denisaepudin@telkomuniversity.ac.id,

³aniqatigi@telkomuniversity.ac.id

Abstrak

Volatilitas merupakan suatu besaran perubahan harga suatu saham yang bertujuan untuk menilai resiko investasi. Volatilitas dapat diartikan bahwa semakin tinggi volatilitas suatu saham maka semakin tinggi pula resiko yang didapat. Peristiwa dimana terjadi pelonjakan maupun penurunan yang sangat drastis pada volatilitas pasar saham dapat disebut sebagai volatilitas tinggi. Sebaliknya, jika terjadi volatilitas yang stabil dan cenderung tidak berubah maka peristiwa itu dapat dikatakan volatilitas rendah. Dalam penelitian ini akan membuktikan bahwa *Financial News*/berita keuangan dapat digunakan untuk memprediksi volatilitas pasar saham dan dibandingkan dengan prediksi volatilitas harga saham penutupan (*Close Price*). Data yang digunakan merupakan data IHSG, untuk *Financial News* diperoleh dari situs Bisnis.com. Sementara, data historical IHSG bersumber dari situs *Yahoo Finance*. Pada *Financial News* diterapkan *Latent Dirichlet Allocation* dengan tujuan untuk memecah dokumen menjadi beberapa bagian kata dan disertai dengan bobot per-token. Kemudian, dilanjutkan dengan penerapan metode Naïve Bayes untuk memprediksi pergerakan volatilitas pasar saham. Serta, menerapkan metode Naïve Bayes untuk memprediksi volatilitas pasar saham dengan melibatkan harga penutupan. Hasil akurasi yang diperoleh dengan menggunakan metode Naïve Bayes dengan melibatkan *Financial News* adalah sebesar 63.6363% untuk data training dan 69.3877% untuk data testing. Sedangkan, hasil akurasi yang diperoleh dengan menggunakan Naïve Bayes melibatkan harga saham penutupan sebesar 53% untuk data training dan 54% untuk data testing.

Kata kunci : volatilitas, financial news, naïve bayes

Abstract

Volatility has meaning of a scale that as a scale of stock market changes which intend to rate the investment risk. Volatility can be interpreted that the higher value of volatility will also make the risk more higher. Volatility where become to very drastic increase or decrease in a stock is called high volatility, otherwise if the volatility turn is stable and tends not to change then it can be said that low volatility. In this research will prove that Financial News / news can be used for predicts stock market volatility and compared with the volatility prediction involving close price. The data used is IHSG data, for the financial news was obtained from the website of Bisnis.com. Meanwhile, the historical data of IHSG was sourced from the Yahoo Finance website. Financial news is implemented by Latent Dirichlet Allocation to split a document into some words which containing tokens / words and weights. Then, applied by the Naïve Bayes method to predict the volatility of the stockmarket. Also, implementing Naïve Bayes to predict stock market volatility by involving close price. Accuracy results obtained using the Naïve Bayes method involving Financial News are 63.6363% for training data and 69.3877% for testing data. While, the accuracy results obtained using Naïve Bayes involving close price are 53% for training data and 54% for data testing.

Keywords: volatility, financial news, naïve bayes

1. Pendahuluan

Latar Belakang

Pada dasarnya harga saham pada sebuah perusahaan bisa berubah menjadi sangat tinggi, cenderung stabil ataupun rendah. Volatilitas merupakan besaran perubahan harga saham dalam kurun waktu tertentu. Ada atau tidaknya suatu volatilitas akan mempengaruhi *return* yang akan di dapatkan oleh para investor. Sehingga, membuat minat para investor menjadi tidak stabil [5].

Volatilitas dapat mempengaruhi investor untuk berinvestasi. Dengan adanya ketidakpastian harga saham akan menyulitkan para investor untuk mengetahui pergerakan harga saham kedepan [6]. Menurut penelitian sebelumnya mengenai *Forecasting daily stock volatility using GARCH model: a comparison between BSE and SSE* (Sasikanta Tripathy, Abdul Rachman. 2013), perkiraan yang tepat terhadap volatilitas pasar saham sangat penting untuk menilai resiko investasi [1]. Pasar yang *volatile* akan menyulitkan perusahaan dalam mengambil langkah untuk menaikkan modalnya dipasaran, maka sangat diperlukan bagi investor untuk mengetahui volatilitas di masa yang akan datang [6]. Berdasarkan paper yang menjadi rujukan penulis mengenai *Financial news predicts stock market volatility better than close price* [2]

mengatakan bahwa, volatilitas bisa diprediksi dengan melibatkan data *Financial News*. Hasil empiris yang ditunjukkan bahwa dengan melibatkan data *Financial News* akurasi yang dihasilkan lebih tinggi dibandingkan dengan prediksi volatilitas dengan melibatkan harga penutupan (*Close Price*) itu sendiri [2].

Permasalahan yang akan dibahas dalam tugas akhir ini adalah bagaimana melakukan prediksi pergerakan volatilitas pasar saham berdasarkan data *Financial News* menggunakan *Latent Dirichlet Allocation* untuk mengambil data. Serta, menerapkan metode Naïve Bayes untuk memprediksi pergerakan volatilitas. Kemudian, hasil prediksi volatilitas yang melibatkan data *Financial News* dibandingkan dengan hasil prediksi volatilitas yang melibatkan data harga penutupan. Data yang digunakan dalam dua perhitungan merupakan indeks saham yang sama yaitu IHSG. Setelah dibandingkan hasil prediksi diantara kedua perhitungan, maka langkah selanjutnya adalah menentukan hasil prediksi volatilitas yang terbaik diantara keduanya.

Topik dan Batasannya

Topik pada Tugas Akhir ini akan memprediksi volatilitas pasar saham dengan menggunakan *Latent Dirichlet Allocation* untuk memperoleh data *Financial News* dan menggunakan metode Naïve Bayes untuk memprediksi pergerakan volatilitas. Setelah itu, mencari hasil terbaik dengan membandingkan hasil prediksi volatilitas yang melibatkan harga penutupan dan hasil prediksi volatilitas yang melibatkan data *Financial News*.

Batasan masalah dalam Tugas Akhir ini diantaranya:

- Data yang digunakan adalah data harga penutupan dan data teks pada *Financial News*
- Data yang diambil berupa data saham IHSG dari tanggal 04 Januari 2017 sampai 6 Juni 2017 untuk data training sedangkan 8 Juni 2017 sampai 25 Agustus 2017 untuk data testing. Sumber data harga penutupan diperoleh dari situs *Yahoo Finance* dan pada data *Financial News* diperoleh dari situs *Bisnis.com*
- Kelas yang digunakan adalah kelas volatilitas naik dan turun

Tujuan

Mengacu pada topik dan batasannya maka tujuan yang akan dicapai dalam pengerjaan Tugas Akhir adalah mengimplementasikan *Latent Dirichlet Allocation* untuk mengambil data pada *Financial News*, dan mengimplementasikan Naïve Bayes untuk memprediksi pergerakan volatilitas harga saham. Serta, membandingkan hasil prediksi volatilitas dengan melibatkan harga penutupan saja.

Organisasi Tulisan

Selanjutnya pada Bab 2 akan dibahas mengenai Studi Terkait, meliputi Volatilitas, *Latent Dirichlet Allocation* dan klasifikasi Naïve Bayes. Pada Bab 3 akan dibahas Sistem yang Dibangun. Pada Bab 4 akan dibahas hasil dan analisis pengujian. Pada Bab 5 akan dibahas mengenai Kesimpulan.

2. Studi Terkait

2.1 Volatilitas

Volatilitas merupakan besaran nilai pelonjakan, penurunan atau pun kestabilan yang terjadi pada pasar saham pada periode tertentu. Volatilitas dihitung dengan menggunakan variansi dari return yang bertujuan untuk memperkirakan resiko dari suatu saham [5]. Adapun untuk mengestimasi volatilitas bisa menggunakan rumus berikut :

$$\sigma^2 \approx \frac{1}{N} \sum_{i=1}^n (r_i - \bar{r})^2 \quad (1)$$

dimana,

$$r_i = \log \left(\frac{S_i}{S_{i-1}} \right) \quad (1.a)$$

r_i = Return dari aset pada waktu i

\bar{r} = Rata-rata return

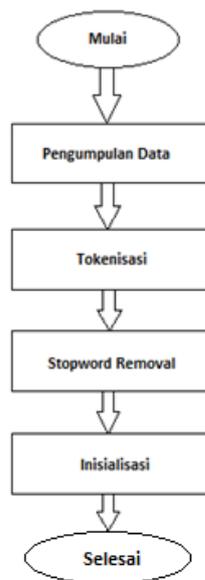
S_i = Harga aset pada waktu i

N = Banyaknya data

Persamaan (1.a) merupakan rumus mencari *return* dimana r_i adalah return dari aset pada waktu i , \bar{r} merupakan *mean* dari *return*, S_i merupakan harga aset pada waktu i sedangkan N merupakan banyaknya data. Dimana hasil dari r_i menjadi input yang digunakan dalam persamaan (1).

2.2 Latent Dirichlet Allocation

Latent Dirichlet Allocation adalah salah satu *Text Modelling* yang digunakan untuk mencari topik pada sebuah dokumen. *Latent Dirichlet Allocation* (LDA) merupakan model dari model Bayes Hirarki yang mengumpulkan beberapa data teks kemudian dibentuk menjadi beberapa topik [4]. LDA berfungsi dalam meringkas, melakukan klusterisasi, menyelesaikan dokumen dalam ukuran yang sangat besar karena LDA menghasilkan daftar topik yang diberi bobot untuk masing-masing dokumen [3]. Adapun distribusi yang digunakan untuk mendapatkan distribusi topik perdokumen disebut distribusi Dirichlet. Kemudian, dalam proses generatif untuk LDA, hasil dari Dirichlet digunakan untuk mengalokasikan kata-kata pada dokumen untuk topik yang berbeda [3]. LDA ini diterapkan pada pemrograman Python menggunakan library *Urllib* dan *Beautifulsoup*, dimana alur LDA dapat dilihat pada gambar di bawah ini :



Gambar 2.2 Alur Metode Latent Dirichlet Allocation

- Pengumpulan Data
Financial News diperoleh dari situs resmi *Bisnis.com* tanggal 4 Januari 2017 sampai dengan tanggal 25 Agustus 2017. Untuk menerapkan LDA penulis menggunakan bahasa pemrograman Python sebagai eksekutor dari penelitian. Modul yang digunakan pada Python yaitu

modul *Urllib* bertujuan untuk mengambil data berita sedangkan *Beautifulsoup* bertujuan untuk mem-*parsing* HTML dan mengekstrak informasi pada berita antara tag tertentu.

- Tokenisasi
Setiap dokumen atau file memiliki sekumpulan kalimat. Tahap tokenisasi ini merupakan proses memecah dokumen teks menjadi beberapa bagian kata [10]. Proses ini dilakukan pada 150 file teks *Financial News* yang telah dikumpulkan.
- Stopword Removal
Pada tahapan ini proses yang dilakukan adalah proses menghilangkan kata-kata yang tidak diperlukan pada file teks *Financial News*, contohnya : ada, yang, telah, setiap, dll.
- Inisialisasi
Proses inisialisasi adalah proses untuk menghitung jumlah frekuensi kemunculan kata pada file teks *Financial News* [10].

2.3 Naïve Bayes

Metode Naïve Bayes memberikan performansi cukup baik untuk banyak kasus modern dengan data yang besar [11]. Algoritmanya menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas [7]. Metode ini memiliki keuntungan yaitu hanya dengan menggunakan data training dalam jumlah yang sedikit untuk bisa mengestimasi parameter pada proses klasifikasi [8]. Klasifikasi Naïve Bayes diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya [9]. Berikut ini merupakan persamaan Naïve Bayes:

$$P(A | B) = \frac{P(B | A) P(A)}{P(B)} \quad (2)$$

B = data dengan *class* yang belum diketahui,

A = hipotesis data merupakan suatu *class* spesifik,

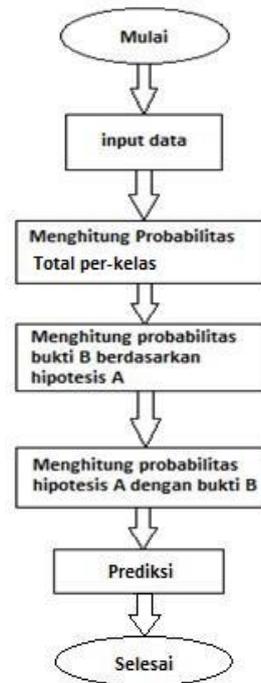
$P(B)$ = probabilitas B ,

$P(A)$ = probabilitas hipotesis A (prior probabilitas),

$P(B|A)$ = probabilitas B berdasarkan kondisi pada hipotesis A ,

$P(A|B)$ = probabilitas hipotesis A berdasar kondisi B (posterior probabilitas).

Berikut ini merupakan alur kerja dari Naïve Bayes :



Gambar 2.3 Alur Metode Naïve Bayes

Persamaan (2) dihitung apabila hanya terdapat satu faktor/fitur. Dalam menjelaskan teorema Naïve Bayes memerlukan beberapa petunjuk untuk menentukan kelas apa yang cocok dengan sampel tersebut. Perhitungan Naïve Bayes dapat melibatkan faktor-faktor yang kompleks. Sehingga, untuk menentukan kelas yang cocok maka diperlukan untuk menghitung peluang kemunculan karakteristik-karakteristik sampel pada kelas tersebut [9]. Jika perhitungan melibatkan jumlah fitur yang banyak, nilai $P(B)$ akan selalu tetap untuk setiap kelas. Oleh karena itu, penjabaran teorema Naïve Bayes dapat disesuaikan sebagai berikut :

$$P(A|B_1, \dots, B_n) = P(A) P(B_1, \dots, B_n|A) \quad (2.a)$$

$$= P(A) P(B_1|A) P(B_2, \dots, B_n|A, B_1) \quad (2.b)$$

$$= P(A) P(B_1|A) P(B_2|A, B_1) P(B_3, \dots, B_n|A, B_1, B_2) \dots P(B_n|A, B_1, B_2, B_{n-1}) \quad (2.c)$$

Persamaan (2.b) dan (2.c) diperoleh dari (2.a) dengan mengasumsikan bahwa B_1, B_2, \dots, B_n saling bebas. Dapat dilihat dari penjabaran tersebut bahwa semakin kompleksnya faktor yang mempengaruhi nilai probabilitas maka perhitungan akan semakin sulit digunakan asumsi independensi tinggi (saling bebas) untuk mempermudah perhitungan yang dilakukan. Sehingga, perhitungan menjadi sebagai berikut :

$$P(A|B_1, \dots, B_n) \approx P(A) \prod_{i=1}^n P(B_i|A) \quad (2.d)$$

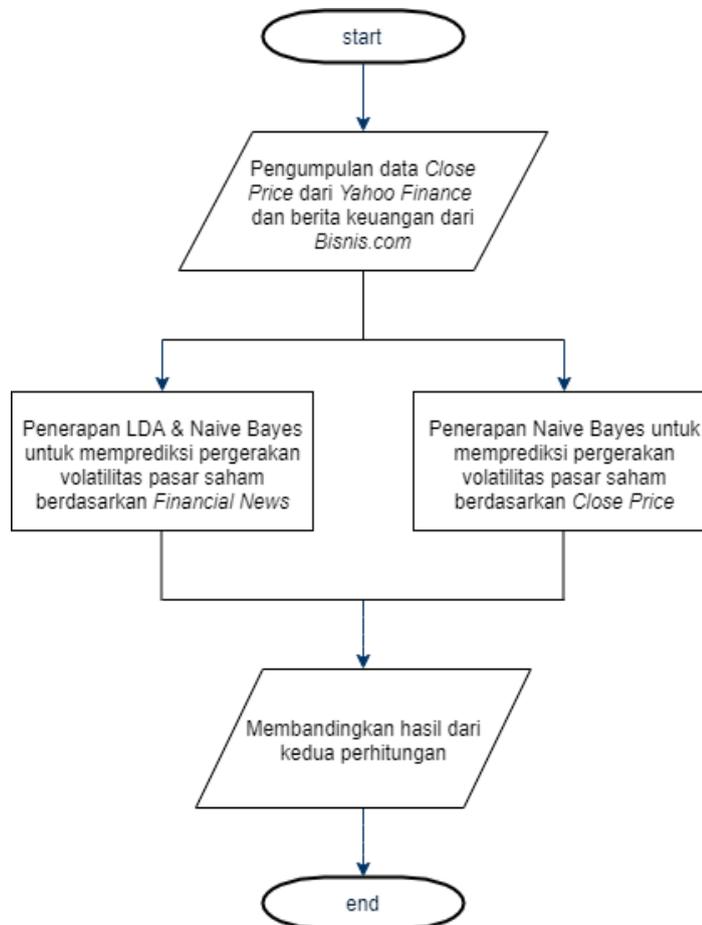
Rumus tersebut kemudian dipilih bagian yang paling besar sebagai kelas hasil prediksi [7]. Nilai akurasi yang digunakan pada penelitian ini adalah sebagai berikut :

$$Akurasi = \frac{\text{Jumlah prediksi yang benar}}{\text{Jumlah prediksi keseluruhan}} \times 100\% \quad (3)$$

3. Sistem yang Dibangun

Berikut rancangan sistem yang dibangun digambarkan pada flowchart sistem pada Gambar 3.1

Gambar 3.1 Flowchart Sistem Prediksi Pergerakan Volatilitas



- Pengumpulan data *Financial News* dan data harga penutupan
Data saham menggunakan *historical* data yang sudah ada sebelumnya di situs resmi *Yahoo Finance* dengan jumlah 150 data. Sedangkan, pada data *Financial News* menggunakan berita yang berasal dari situs *Bisnis.com* bagian ekonomi dengan jumlah 150 berita. Kelas yang digunakan ada dua yaitu volatilitas tinggi dan rendah. Untuk volatilitas naik dikonversikan menjadi “1”, sedangkan untuk volatilitas turun dikonversikan menjadi “0”
- Penerapan LDA
Pada setiap dokumen terdapat beberapa topik yang berbeda. Cara kerja LDA ini adalah mengklasifikasikan sebuah dokumen yang menggambarkan suatu topik tertentu dengan kata-kata yang spesifik dalam topik tersebut. LDA ini diterapkan melalui bahasa pemrograman Python dengan mengekstrak data pada *Financial News*

- Penerapan metode Naïve Bayes
 1. Penerapan Naïve Bayes pada *Financial News*

Output dari Latent Dirichlet Allocation pada *Financial News* yang telah didapat, diterapkan pada metode Naïve Bayes untuk memperoleh prediksi volatilitas. Persamaan yang digunakan adalah persamaan (2.d) karena atribut yang digunakan untuk perhitungan probabilitas berjumlah banyak, sehingga untuk memudahkan perhitungan maka menggunakan persamaan tersebut. Sedangkan pada data harga penutupan, persamaan yang digunakan adalah persamaan (2). Dalam data harga penutupan ini terjadi proses diskritisasi, dimana harga *Close Price* sebelumnya dibandingkan dengan harga *Close Price* setelahnya. Hasil dari proses tersebut dijadikan sebagai input dalam penerapan metode Naïve Bayes
 2. Penerapan Naïve Bayes pada *Close Price*
 - a. Membandingkan nilai *Close Price* sebelumnya dengan setelahnya

Melakukan perbandingan antara nilai *Close Price* sebelumnya dengan nilai *Close Price* setelahnya. Apabila nilai *Close Price* setelahnya lebih besar daripada nilai *Close Price* sebelumnya maka dikonversikan ke "1". Sebaliknya, jika nilai *Close Price* setelahnya lebih rendah daripada nilai *Close Price* sebelumnya maka dikonversikan ke "0"
 - b. Menghitung variansi atau estimasi volatilitas *Close Price*

Melakukan perhitungan variansi atau estimasi volatilitas dengan melibatkan nilai return
 - c. Membandingkan nilai volatilitas sebelumnya dengan setelahnya

Setelah mendapatkan nilai variansi, kemudian melakukan perbandingan antara nilai variansi setelahnya dengan nilai variansi sebelumnya. Jika nilai variansi setelahnya lebih besar dibandingkan dengan nilai variansi sebelumnya maka dikonversikan ke "1". Sebaliknya, jika nilai variansi setelahnya lebih rendah daripada nilai variansi sebelumnya maka dikonversikan ke "0"
- Perbandingan kedua perhitungan

Pada bagian ini, hasil prediksi volatilitas yang melibatkan data *Financial News* akan dibandingkan dengan hasil prediksi volatilitas yang melibatkan harga penutupan.

4. Evaluasi

4.1 Hasil Pengujian

4.1.1 Latent Dirichlet Allocation

Pada *Financial News* diterapkan metode LDA, dimana metode tersebut mengeluarkan output berupa beberapa token/kata yang mendominasi, serta mengeluarkan output berupa nilai bobot dari kata tersebut untuk dikelompokkan menjadi topik yang sama. Modul yang digunakan berupa *Urllib* dan *Beautifulsoup* untuk menarik data *Financial News* yang ada pada situs resmi *Bisnis.com*. Dengan menginput jumlah topik=1, jumlah kata=20 dengan jumlah berita yang digunakan adalah 150 berita, maka diperoleh hasil sebagai berikut :

```

C:\Program Files (x86)\Python37-32\python.exe "D:/Nadhieya/TA Bismillah/data TA/data yang digunakan/parse.py"
[['ihsg']]
['hari']
['perdagangan']
['wib']
['saham']
['indeks']
['harga']
['level']
['poin']
['menguat']
['agustus']
['pergerakan']
['gabungan']
['bursa']
['rekomendasi']
['emas']
['ditutup']
['naik']
['ago']
['melemah']]
[[0.016 0. 0.073 0.035 0.028 0. 0.028 0.045 0.029 0.017 0. 0.039
0.03 0.045 0. 0. 0.038 0.031 0.062 0. 0. 0. 0. 0.034
0. 0. 0. 0. 0. 0. 0.024 0. 0.022 0. 0. 0.112
0.091 0. 0. 0. 0.036 0.025 0.039 0. 0. 0. 0. 0.
0. 0.091 0.098 0.06 0. 0.076 0.082 0. 0. 0. 0.023 0.
0. 0. 0.034 0. 0.055 0.032 0. 0. 0.041 0.042 0.105 0.
0. 0.06 0.035 0.075 0. 0.031 0.035 0.03 0. 0.06 0.05 0.113
0.054 0. 0.026 0. 0. 0. 0. 0. 0. 0.049 0. 0.
0. 0.071 0. 0. ]
[0. 0.027 0. 0.021 0. 0. 0.02 0.037 0.055 0. 0. 0.
0. 0.021 0. 0. 0.068 0.047 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0.021 0. 0. 0.018 0. 0. 0. 0.

```

Gambar 4.1 Hasil running Latent Dirichlet Allocation di Python

Nilai bobot yang dihasilkan dari penerapan LDA diatas diinisialisasikan menjadi “1”, jika terdapat token yang muncul dalam teks berita dan dikatakan “0”, jika token tidak muncul dalam teks berita. Berdasarkan gambar di atas terdapat hasil dari Latent Dirichlet Allocation dengan tujuan untuk mendapatkan beberapa token disertai dengan bobot setiap token per berita. Output dari LDA dituliskan ke excel, pada excel terjadi perhitungan secara manual untuk mendapatkan target/kelas. Setelah didapatkan data *Financial News* beserta target, data diimplementasikan pada metode Naïve Bayes untuk memprediksi pergerakan volatilitas pasar saham.

4.1.2 Naïve Bayes

4.1.2.1 Implementasi Naïve Bayes menggunakan Data *Financial News*

Data berita yang didapatkan melalui Latent Dirichlet Allocation kemudian dilakukan prediksi pergerakan volatilitas dengan menggunakan Naïve Bayes. Dengan data training dan juga data testing berukuran sama yaitu sebanyak 99 baris dan 20 kolom. Perhitungan yang dilakukan menggunakan persamaan (2.d) karena faktor-faktor syarat berjumlah banyak. Semakin banyak faktor penentu dan semakin kompleks syarat akan mempengaruhi nilai probabilitas sehingga menyebabkan perhitungan menjadi lebih sulit. Oleh karena itu, perhitungan digunakan menggunakan persamaan (2.d). Hasil prediksi volatilitas yang didapatkan dari *Financial News* adalah sebagai berikut :

Tabel 1. Hasil Perhitungan Data *Financial News*

Data	Banyak berita	Akurasi
Training	99	63.6363%
Testing	49	69.3877%

4.1.2.2 Implementasi Naïve Bayes menggunakan Data Harga Penutupan

Untuk mendiskritisasi data maka nilai harga penutupan sebelumnya dibandingkan nilai harga penutupan setelahnya yang kemudian akan dikonversikan ke “0” jika harga penutupan turun dan dikonversikan ke “1” jika harga penutupan naik. Hasil dari perbandingan tersebut digunakan sebagai input yang akan diterapkan pada Naïve Bayes.

Untuk mendapatkan standar deviasi dari harga penutupan, maka digunakan rumus standar deviasi. Standar deviasi yang didapat berdasarkan jumlah dari 10 data sebelumnya. Target pada data penutupan ditentukan apabila nilai dari standar deviasi setelahnya lebih besar dari standar deviasi sebelumnya maka dikonversikan ke “1”, begitu sebaliknya. Setelah mendapat input dan target, maka diprediksi dengan persamaan (2). Hasil prediksi volatilitas melibatkan harga penutupan adalah sebagai berikut :

Tabel 2. Hasil Perhitungan Data Harga Penutupan

Data	Banyak berita	Akurasi
Training	99	53%
Testing	49	54%

4.1.2.3 Perbandingan Hasil Implementasi Naïve Bayes

Berikut merupakan perbandingan hasil implementasi Naïve Bayes berdasarkan *Financial News* dan harga penutupan :

Tabel 3. Perbandingan Hasil Akurasi *Financial News* dan Harga Penutupan

Data	Akurasi	
	Financial News	Harga penutupan
Training	63.6363%	53%
Testing	69.3877%	54%

Pada tabel diatas menjelaskan bahwa akurasi yang dihasilkan dengan melibatkan *Financial News* lebih besar yaitu sebesar 63.6363% pada data training dan akurasi sebesar 69.3877% pada data testing. Sedangkan, akurasi yang dihasilkan dengan melibatkan harga penutupan adalah sebesar 53% pada data training dan akurasi sebesar 54% pada data testing.

4.2 Analisis Hasil Pengujian

Setelah mendapatkan hasil prediksi volatilitas dengan melibatkan data *Financial News* dan dibandingkan dengan prediksi volatilitas yang melibatkan harga penutupan, maka selanjutnya melakukan analisis hasil pengujian. Data *Financial News* yang diinput melibatkan informasi yang komprehensif dari suatu berita, kemudian, data diimplementasikan menggunakan rumus penjabaran Naïve Bayes. Sehingga, dalam perhitungannya terdapat informasi yang relevan dengan naik turun harga saham, sedangkan pada data harga penutupan hanya menginput satu atribut untuk diimplementasikan ke rumus Naïve Bayes maka diperlukan hasil yang lebih baik. Berikut ini merupakan hasil analisis dari data *Financial News* apabila diinputkan jumlah atribut yang berbeda. Hasil yang diperoleh adalah sebagai berikut:

Tabel 4. Hasil Perhitungan Akurasi dengan Atribut Berbeda pada Data Training

Banyak berita	Jumlah atribut	Akurasi
99	20	63.6363%
99	15	58.5858%
99	10	54.5454%

Tabel 5. Hasil Perhitungan Akurasi dengan Atribut Berbeda pada Data Testing

Banyak berita	Jumlah atribut	Akurasi
49	20	69.3877%
49	15	75.5102%
49	10	61.2245%

Tabel 4 dan tabel 5 menunjukkan bahwa jumlah atribut tidak berpengaruh secara signifikan terhadap tingkat akurasi, namun akurasi memiliki kecenderungan naik dengan bertambahnya jumlah atribut. Tingkat akurasi pada data training dengan 20 atribut mencapai 63.6363%, pada 15 atribut menghasilkan akurasi sebesar 58.5858% dan sebesar 54.5454% dengan 10 atribut. Pada data testing dengan 20 atribut menghasilkan tingkat akurasi sebesar 69.3877%, pada 15 atribut menghasilkan akurasi sebesar 75.5102% dan pada 10 atribut menghasilkan akurasi sebesar 61.2245%.

5. Kesimpulan

Berdasarkan hasil perhitungan pada Tugas Akhir ini dapat disimpulkan bahwa :

1. Akurasi data training yang diperoleh dengan melibatkan *Financial News* adalah sebesar 63.6363% dan akurasi data testing adalah sebesar 69.3877%. Sedangkan, akurasi data training yang diperoleh dengan melibatkan *Close Price* adalah sebesar 53% dan akurasi data testing adalah sebesar 54%. Berdasarkan hal tersebut, dapat disimpulkan bahwa hasil akurasi dengan melibatkan *Financial News* lebih baik daripada hasil akurasi yang diperoleh dengan melibatkan *Close Price*.
2. Akurasi data training sebanyak 20 atribut adalah sebesar 63.6363%, sebanyak 15 atribut adalah sebesar 58.5858% dan sebanyak 10 atribut adalah sebesar 54.5454%. Pada data testing tingkat akurasi yang dihasilkan 20 atribut adalah sebesar 69.3877%, sebanyak 15 atribut adalah sebesar 75.5102% dan sebanyak 10 atribut adalah sebesar 61.2245%.

Daftar Pustaka

- [1] Sasikanta Tripathy, Abdul Rahman. *Forecasting Daily Stock Volatility Using GARCH Model:*

- A Comparison between BSE and SSE, The UIP Journal of Applied Finance*, Vol. 19, No. 4, 2013.
- [2] Adam Atkins, Mahesan Niranjana, Enrico Gerding. *Financial News Predicts Stock Market Volatility Better than Close Price*, *The Journal of Finance and Data Science*, Vol. 4, Issue 2, Pages 120-137, June 2018.
- [3] I Made Kusnanta Bramantya Putra, Renny Pradina Kusumawardani. Analisis Topik Informasi Publik Media Sosial di Surabaya Menggunakan Pemodelan Latent Dirichlet Allocation (LDA), *Jurnal Teknik ITS*, Vol. 6, No. 2, Institut Teknologi Sepuluh November, 2017.
- [4] David M. Blei, Andrew Y. Ng, Michael I. Jordan. *Alokasi Dirichlet laten*, Volume 3, *Journal of Machine Learning*, 2003.
- [5] Linda Karlina Sari, Noer Azam Achsanib, Bagus Sartono. *Modelling Volatility of Return Stock Index: Evidence from Asian Countries*, *Jurnal Ekonomi dan Pembangunan Indonesia*, Vol. 18 No. 1, Juli 2017.
- [6] Andi Kartika. Volatilitas Harga Saham di Indonesia dan Malaysia, *Jurnal Ilmu Ekonomi Aset*, ISSN 1693- 928X, Vol. 12, No. 1, Hal. 17-26, Maret 2010.
- [7] Yani Parti Astuti, Usman Sudibyo, Achmad Wahid Kurniawan, Yuniarsi Rahayu. Algoritma Naive Bayes dengan Fitur Seleksi untuk mengetahui Hubungan Variabel Nilai dan Latar Belakang Pendidikan, Vol. 9, No. 1, April 2018.
- [8] Pattekari, S. A., Parveen, A. *Prediction System for Heart Disease Using Naive Bayes*, *International Journal of Advanced Computer and Mathematical Sciences*, ISSN 2230-9624, Vol. 3, No. 3, Hal. 290-294, 2012.
- [9] Bustami. Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi, *TECHSI : Jurnal Penelitian Teknik Informatika*, Vol. 3, No.2, Hal. 127-146, 2013.
- [10] Putu Manik Prihatini, I Ketut Suryawan, I Nyoman Mandia. Metode Latent Dirichlet Allocation untuk Ekstraksi Dokumen, *Jurnal Logic*, Vol. 17, No. 13, November 2017.
- [11] Suyanto. 2018. *Machine Learning Tingkat Dasar dan Lanjut*. Bandung : Informatika Bandung.