



## Analisis Data Produksi Biskuit Dengan Algoritma Naive Bayes Dan Random Forest

\* Sabaruddin <sup>1</sup>, Agung Budi Santoso <sup>2</sup>, Sajarwo Anggai <sup>3</sup>

<sup>1,2,3</sup> Teknik Informatika Program Pascasarjana, Universitas Pamulang, Kota Tangerang Selatan, Banten

Email: <sup>1</sup> sabaruddin02@gmail.com, <sup>2</sup> dosen02680@unpam.ac.id, <sup>3</sup> sajarwo@gmail.com

### ABSTRACT

*In the manufacturing industry, production problems often occur, often production does not match market demand, production is not well planned, therefore this study aims to develop a classification model using machine learning based on the Naive Bayes and Random Forest algorithms to classify biscuit production data. The main focus of this study is to utilize variables such as dough, number of mixers, production time parameters, and other relevant production factors to improve accuracy in classification. The dataset used in this study includes information from several previous production periods, namely data in 2019-2023, which is then used to train and test the Naive Bayes and Random Forest algorithm models. The training and validation process is carried out using commonly used model performance evaluation techniques. The results of the study show that the Random Forest model is able to provide high accuracy, namely 97.54% while Naive Bayes is 96.45%. Further analysis was also carried out to identify the variables that most influence production results, providing additional insights for optimizing the production process. The results of this study can contribute to the development of classification models for the food and beverage industry, especially in biscuit products, but also offer a more specific view of the factors that influence biscuit production. The implementation of this study can be a basis for manufacturers to make more precise and effective decisions in managing their production.*

**Keywords:** Classification, Naive Bayes, Random Forest, Production

### ABSTRAK

Dalam industri manufaktur sering terjadi permasalahan produksi, seringkali produksi tidak sesuai dengan permintaan pasar, produksi tidak terplaning dengan baik oleh karena itu penelitian ini bertujuan untuk mengembangkan model klasifikasi dengan menggunakan machine learning berbasis algoritma Naive Bayes dan Random Forest untuk mengklasifikasi data produksi biskuit. Fokus utama penelitian ini adalah memanfaatkan variabel-variabel seperti adonan, jumlah mixer, parameter waktu produksi, dan faktor-faktor produksi lainnya yang relevan untuk meningkatkan akurasi dalam mengklasifikasi. Dataset yang digunakan dalam penelitian ini mencakup informasi dari beberapa periode produksi sebelumnya yaitu data pada tahun 2019-2023, yang kemudian digunakan untuk melatih dan menguji model algoritma Naive Bayes dan Random Forest. Proses pelatihan dan validasi dilakukan dengan memanfaatkan teknik-teknik evaluasi kinerja model yang umum digunakan. Hasil penelitian menunjukkan bahwa model Random Forest mampu memberikan akurasi yang tinggi yaitu 97,54% sedangkan Naive Bayes 96,45%. Analisis lanjutan juga dilakukan untuk mengidentifikasi variabel yang paling berpengaruh terhadap hasil produksi, memberikan wawasan tambahan untuk optimalisasi proses produksi. Hasil penelitian ini dapat memberikan kontribusi pada pengembangan model klasifikasi untuk industri makanan dan minuman, khususnya pada produk biskuit, tetapi juga menawarkan pandangan yang lebih spesifik terkait faktor-faktor yang memengaruhi

produksi biskuit. Implementasi penelitian ini dapat menjadi landasan bagi produsen untuk mengambil keputusan yang lebih tepat dan efektif dalam mengelola produksi mereka.

**Kata Kunci** : Klasifikasi, *Naïve Bayes*, *Random Forest*, Produksi

## 1. PENDAHULUAN

Dalam beberapa tahun terakhir, metode-metode berbasis teknologi informasi, seperti *Machine Learning* (ML), telah menjadi alat yang efektif untuk memprediksi berbagai parameter dalam berbagai industri. Algoritma *Naïve Bayes* dan *Random Forest* merupakan salah satu algoritma *Machine Learning* (ML) yang umum digunakan untuk memodelkan hubungan antara variabel independen dan dependen. Dengan menerapkan algoritma ini pada data produksi biskuit, dapat dibangun model prediktif yang dapat membantu industri dalam merencanakan produksi mereka secara lebih akurat.

Adanya fluktuasi dalam permintaan pasar, perubahan musiman, dan faktor-faktor eksternal lainnya membuat prediksi dan klasifikasi produksi menjadi suatu tantangan yang kompleks. Oleh karena itu, keberhasilan implementasi model *Machine Learning*, terutama dengan menggunakan algoritma *Naïve Bayes* dan *Random Forest* dapat memberikan manfaat signifikan bagi industri biskuit dalam meningkatkan efisiensi produksi, mengoptimalkan persediaan, dan mengurangi risiko ketidaksesuaian antara produksi dan permintaan.

Penelitian ini bertujuan untuk menggali potensi dan efektivitas penerapan *Machine Learning* dengan algoritma *Naïve Bayes* dan *Random Forest* dalam meramalkan dan mengklasifikasi jumlah produksi biskuit. Dengan memahami tren produksi, pola musiman, dan faktor-faktor lain yang memengaruhi produksi biskuit, diharapkan dapat diciptakan model prediktif dan klasifikasi yang dapat membantu industri untuk merencanakan produksi dengan lebih tepat dan efisien. Selain itu, hasil penelitian ini juga dapat memberikan kontribusi pada perkembangan pengetahuan dalam bidang *Machine Learning* dan aplikasinya dalam industri makanan.

## 2. METODE

### 2.1. Metode Penelitian

Penelitian ini mengungkap pendekatan deskriptif kuantitatif dan kajian literatur. Disebut sebagai deskriptif kuantitatif karena data yang diakses berupa data numerik yang telah ada (data numerik), kemudian dianalisis dan disusun sesuai kebutuhan dan

ketentuan penelitian. Proses deskripsi data dilakukan terlebih dahulu sebelum melakukan analisis data dengan tujuan untuk mendapatkan gambaran umum dari setiap variabel data.

## **2.2. Data Dan Sumber Data**

Proses pengumpulan data dilakukan secara menyeluruh selama periode dari bulan Januari hingga bulan Desember selama lima tahun yaitu dari tahun 2019-2023, sesuai dengan rentang waktu yang telah ditentukan untuk menggambarkan perkembangan dan variasi data yang signifikan.

## **2.3. Metoda Pengumpulan Data**

Dalam penelitian ini, peneliti memperoleh data dari laporan hasil produksi biskuit sebagai salah satu referensi utama.

Metode pengumpulan data yang digunakan melibatkan dua aspek penting, yaitu:

### 1) Observasi

Observasi merupakan teknik pengumpulan data yang dilakukan secara cermat oleh peneliti. Hasil observasi berupa data angka mengenai jumlah hasil produksi, memungkinkan peneliti untuk mendapatkan gambaran yang lebih nyata terhadap peristiwa atau kejadian yang menjadi fokus penelitian. Observasi ini bertujuan untuk memberikan jawaban konkret terhadap pertanyaan penelitian.

### 2) Dokumen

Selain melalui observasi, data juga diperoleh melalui fakta-fakta yang terdokumentasi dalam bentuk dokumen. Informasi-informasi ini dicatat secara manual dan dijadikan tambahan untuk memperkaya konteks penelitian. Pendekatan ini memastikan keberagaman dan kelengkapan data yang diperlukan untuk mendukung analisis dalam penelitian ini.

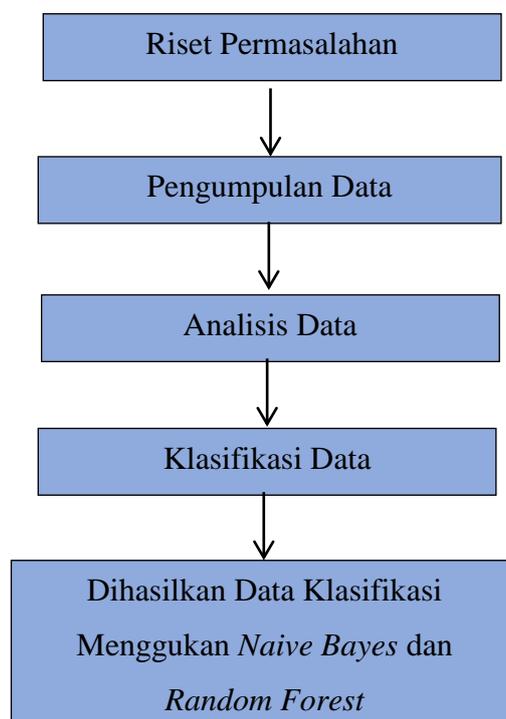
## **2.4. Teknik Analisis**

Analisis data dilaksanakan setelah peneliti menyelesaikan tahap pengumpulan data. Proses ini mencakup pencarian dan penyusunan data secara sistematis, yang diperoleh melalui catatan lapangan, observasi, dan dokumentasi. Data disusun ke dalam kategori, dijabarkan menjadi unit-unit data, dan dianalisis untuk membuat kesimpulan yang dapat dipahami dengan mudah oleh peneliti sendiri maupun pihak lain yang berkepentingan. Parameter utama dalam penelitian ini melibatkan informasi Adonan,

jumlah *Mixer*, *Cream*, jumlah *Broken* dan lain-lain. Selanjutnya, semua data yang terkumpul akan digunakan untuk melakukan prediksi menggunakan algoritma *Naive Bayes* dan *Random Forest*.

## 2.5. Langkah-Langkah Penelitian

Penelitian ini bertujuan untuk mengembangkan model prediksi produksi biskuit menggunakan algoritma *Naive Bayes* dan *Random Forest*. Dengan merinci faktor-faktor produksi dan menerapkan teknik *sampling*, penelitian ini bertujuan untuk meningkatkan akurasi prediksi dan memberikan wawasan yang berharga bagi industri makanan.



**Gambar 1** Langkah-Langkah Penelitian

Keterangan :

Peneliti memulai penelitian dengan menganalisis data produksi biskuit, yang diikuti dengan tahap pengumpulan data. Sebelum data diproses, peneliti menetapkan tujuan penelitian dan melakukan studi pustaka yang sesuai dengan topik sebagai referensi. Setelah itu, dilakukan pendekatan pengolahan data menggunakan algoritma *Naive Bayes* dan *Random Forest*. Proses ini diikuti dengan pengujian dan percobaan *Naive Bayes* dan *Random Forest*. Pada tahap analisis hasil, peneliti memperoleh hasil dari penerapan metode *Naive Bayes* dan *Random Forest*.

### 3. HASIL DAN PEMBAHASAN

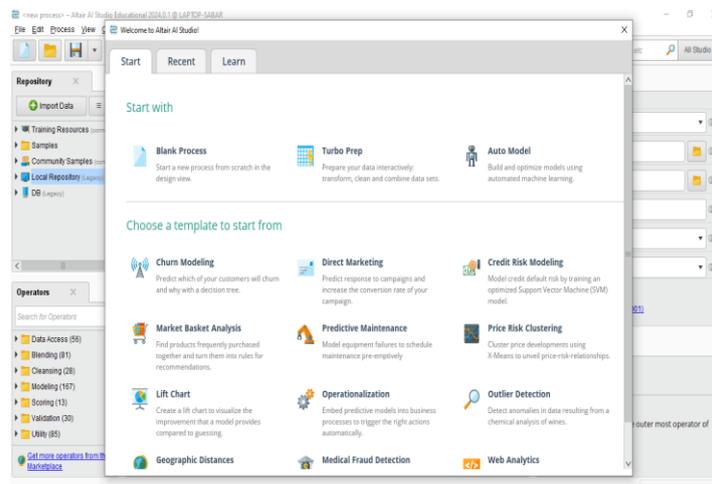
Pada penelitian ini, penulis menggunakan dataset yang diambil dari melalui akses ke server lokal yang tersedia bagi semua pengguna internal perusahaan. Dari dataset tersebut akan diambil sampel data pada tahun 2019-2023 dari bulan januari sampai bulan desember. Data yang akan diambil adalah data produksi biskuit tahun 2019-2023. Variabel yang digunakan adalah Jumlah Total *Broken*. Data tersebut kemudian akan diolah menggunakan software *Rapidminer Studio 10.3* dengan pengujian menggunakan beberapa algoritma yaitu algoritma *Naïve Bayes* dan *Random Forest*.

#### 3.1. Data Preparation

Dataset yang akan diolah ke aplikasi *Rapidminer* bisa menggunakan format Excel, *Uniform Resource Locator (URL)*, *Statistical Product and Service Solutions (SPSS)*, *Statistika* dan *Data (Stata)*, *Attribute-Relation File Format (ARFF)*, *eXtensible attribute-Relation File Format (XRFF)*, format file yang biasanya digunakan oleh perangkat lunak *database (Dbase)*, *C4.5*, *BibTex*, *DasyLab*, *XML*, *Access* dan *Command Separated Values (CSV)*. *Dataset* yang digunakan terdiri dari *feature Date*, No Urut, No Urut2, Nomor *Oven*, No, Kode *Oven*, Kode, *Item*, Adonan, Jml *Mixer*, *Start*, *Finish*, Jam, Menit, *Evaporation*, % *Wash*, % *Sugar*, % *Oil*, % *Salt*, % *Cream*, Biskuit -, Cr Biskuit +, Cr *Broken* -, Cr *Broken* +, Cr *Broken* kulit, % *Broken*, Total *Broken*, Klasifikasi, Total *Product*.

Langkah-langkah *preprocessing* data adalah sebagai berikut :

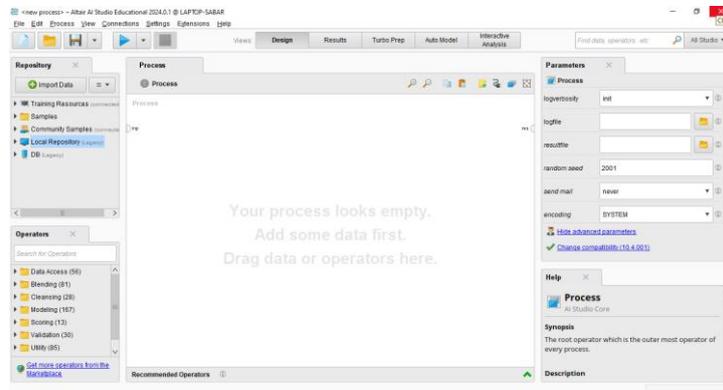
- Membuka dan menjalankan aplikasi *Rapidminer*



Gambar 4.1 Tampilan Awal *Rapidminer*

Pada Gambar 4.1 adalah tampilan dari tools *Rapidminer 10.3* yang dijadikan sebagai *tools* untuk menguji *dataset* Produksi biskuit tahun 2019-2023 menggunakan *cross validation*.

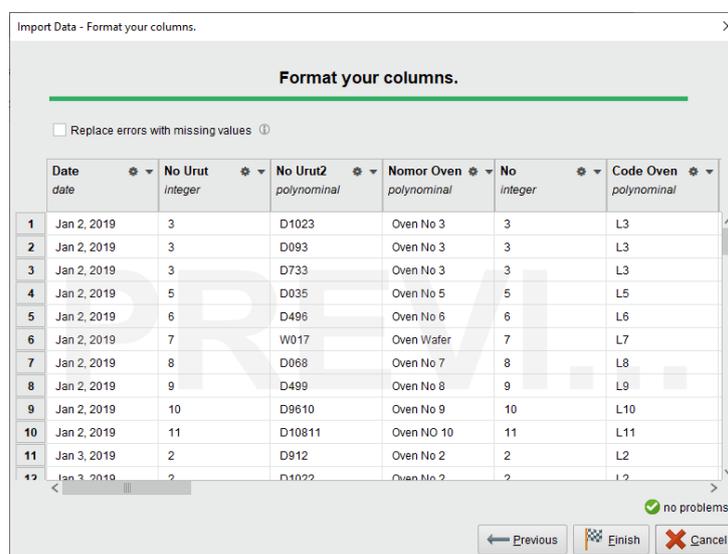
- Klik tombol *Blank Process* untuk membuka antarmuka *Rapidminer*



**Gambar 4.2** Tampilan Antarmuka *Rapidminer*

Dari Gambar 4.3 diatas bisa dilihat fungsi-fungsi dari aplikasi *Rapidminer* ,dibagian tampilan anatarmuka ada *Repository*, *Process*, *Parameters*, *Operators* dan *Help*.

- Buka file *dataset* yang akan diuji

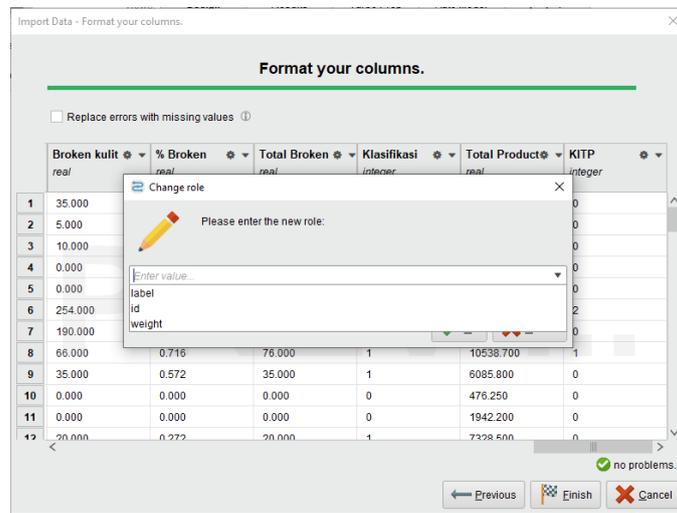


	Date	No Urut	No Urut2	Nomor Oven	No	Code Oven
	date	integer	polynomial	polynomial	integer	polynomial
1	Jan 2, 2019	3	D1023	Oven No 3	3	L3
2	Jan 2, 2019	3	D093	Oven No 3	3	L3
3	Jan 2, 2019	3	D733	Oven No 3	3	L3
4	Jan 2, 2019	5	D035	Oven No 5	5	L5
5	Jan 2, 2019	6	D496	Oven No 6	6	L6
6	Jan 2, 2019	7	W017	Oven Wafer	7	L7
7	Jan 2, 2019	8	D068	Oven No 7	8	L8
8	Jan 2, 2019	9	D499	Oven No 8	9	L9
9	Jan 2, 2019	10	D9610	Oven No 9	10	L10
10	Jan 2, 2019	11	D10811	Oven NO 10	11	L11
11	Jan 3, 2019	2	D912	Oven No 2	2	L2
12	Jan 3, 2019	2	D1022	Oven No 2	2	L2

**Gambar 4.3** Tampilan *dataset* Produksi Biskuit Tahun 2019-2023

Tampilan dari Gambar 4.4 *dataset* menunjukkan karakteristik sebuah *dataset* yang nantinya penulis akan mengarahkan *dataset* ke penelitian klasifikasi sesuai dari karakteristik *dataset*. Dari tampilan diatas bisa diketahui terdapat jumlah *attribute/features* = 29, jumlah *intances* = 28322 dan jumlah *Missing values* (nilai yang hilang pada setiap attribute dan class) = 0.

Penentuan label dari dataset Produksi Biskuit Tahun 2019-2023 dilakukan di *processing* dengan cara klik → *edit* maka akan tampil seperti dibawah ini :



**Gambar 4.4** Pemilihan *Label*

Pada Gambar 4.5 cara pemilihan *Label* sebuah dataset di *tools Rapidminer*, feature yang akan dijadikan *Label* adalah *Klasifikasi*.

Selanjutnya dataset di *normalize* untuk memodifikasi nilai dalam *variable* sehingga kita dapat mengukurnya dalam skala umum. Kemudian melakukan konversi tipe data untuk label dari *numerik* ke *nominal*.

**3.2. Pengujian Model**

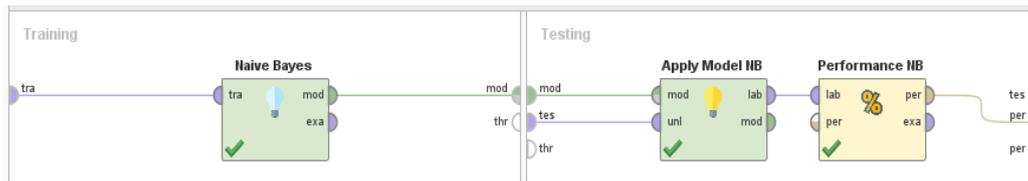
Pada tahap ini dilakukan desain model yang akan digunakan untuk melakukan pengujian terhadap data *training* dan *testing*. Pengujian terhadap data *testing* dan training akan dilakukan menggunakan dua algoritma yang memiliki akurasi tertinggi dari keseluruhan algoritma yang digunakan pada penelitian terdahulu berdasarkan pembahasan pada bab sebelumnya. Algoritma yang digunakan adalah algoritma *Naïve Bayes* dan *Random Forest*.

**Tabel 4.1** Model yang akan diujikan

Model	Algoritma	Label
Model 1	<i>Naïve Bayes</i>	Total Broken
Model 2	<i>Random Forest</i>	Total Broken

Pada Tabel 4.1 dari model diatas penulis akan menguji dua algoritma yaitu *Naïve Bayes* dan *Random Forest* dengan label *Total Broken*. Dari lima algoritma ini akan dicari model yang terbaik.

### 1. Naïve Bayes



**Gambar 4.5** Naive Bayes pada Rapidminer

Pengujian model yang pertama adalah Model 1 yaitu algoritma *Naïve Bayes* dengan *Label Total Broken*. Pengujian dilakukan menggunakan aplikasi *RapidMiner Studio 10.3*, menggunakan skema desain model seperti berikut ini :

- a. Pada tahap pertama dilakukan import data dari operator *Read Exce* dengan nama dataset “*General Report 2019-2023.xlsx*”.
- b. Tahap selanjutnya adalah melakukan *splitting dataset* menjadi 2 bagian, yaitu *Training* dan *Testing* yang diporsi sebanyak 0,8:0,2.
- c. Selanjutnya adalah memproses data yang sudah dibagi tersebut menggunakan algoritma *Naïve Bayes* dengan *Dataset General Report 2019-2023.xlsx*.
- d. Lalu dilakukan validation terhadap data dan diuji menggunakan algoritma *Naïve Bayes* dengan *Label Total Broken* sebelum dilakukan pengukuran performa.
- e. Kemudian dilakukan pengukuran performa dari data tersebut, menggunakan parameter evaluasi performa *Accuracy* dan *Confusion Matrix*.

accuracy: 96.45% +/- 0.44% (micro average: 96.45%)

	true 1	true 0	true 2	class precision
pred. 1	14702	5	10	99.90%
pred. 0	0	3526	0	100.00%
pred. 2	685	3	894	56.51%
class recall	95.55%	99.77%	98.89%	

**Gambar 4.6** Accuracy pada model satu

Pada gambar 4.7 dapat dilihat tingka *accuracy* algoritma *Naïve Bayes* pada dataset *General Report 2019-2023.xlsx* mencapai **96.45%**

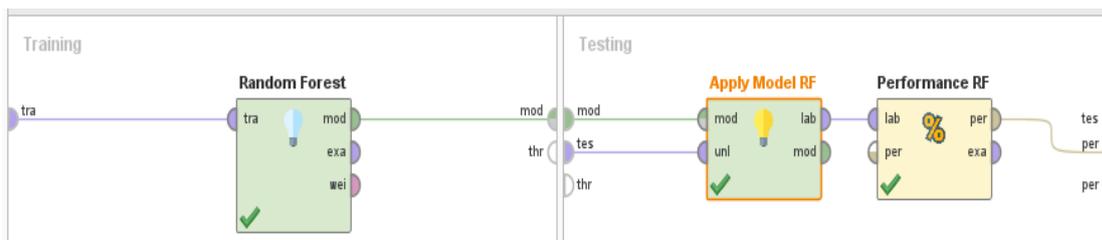
## PerformanceVector

```

PerformanceVector:
accuracy: 96.45% +/- 0.44% (micro average: 96.45%)
ConfusionMatrix:
True:   1      0      2
1:     14702   5      10
0:       0    3526     0
2:       685     3    894
  
```

**Gambar 4.7** *Accuracy* pada model satu

## 2. *Random Forest*



**Gambar 4.8** *Random Forest* pada *Rapidminer*

Pengujian model yang ke dua adalah Model 2 yaitu algoritma *Random Forest* dengan *Label Total Broken*. Pengujian dilakukan menggunakan aplikasi *RapidMiner Studio 10.3* menggunakan skema desain model seperti berikut ini :

- Pada tahap pertama dilakukan import data dari operator *Read Exce* dengan nama dataset "*General Report 2019-2023.xlsx*".
- Tahap selanjutnya adalah melakukan *splitting dataset* menjadi 2 bagian, yaitu *Training* dan *Testing* yang diporsi sebanyak 0,8:0,2.
- Selanjutnya adalah memproses data yang sudah dibagi tersebut menggunakan algoritma *Random Forest* dengan Dataset *General Report 2019-2023.xlsx*.
- Lalu dilakukan validation terhadap data dan diuji menggunakan algoritma *Random Forest* dengan *Label Total Broken* sebelum dilakukan pengukuran performa.
- Kemudian dilakukan pengukuran performa dari data tersebut, menggunakan parameter evaluasi perfoma *Accurasy* dan *Confusion Matrix*.

accuracy: 97.34% +/- 0.42% (micro average: 97.34%)

	true 1	true 0	true 2	class precision
pred. 1	15277	389	28	97.34%
pred. 0	55	3145	0	98.28%
pred. 2	55	0	876	94.09%
class recall	99.29%	88.99%	96.90%	

**Gambar 4.9** Accuracy Pada Model 2

Pada gambar 4.10 dapat dilihat tingka *accuracy* algoritma *Random Forest* pada dataset *General Report 2019-2023* mencapai **97.34%**.

## PerformanceVector

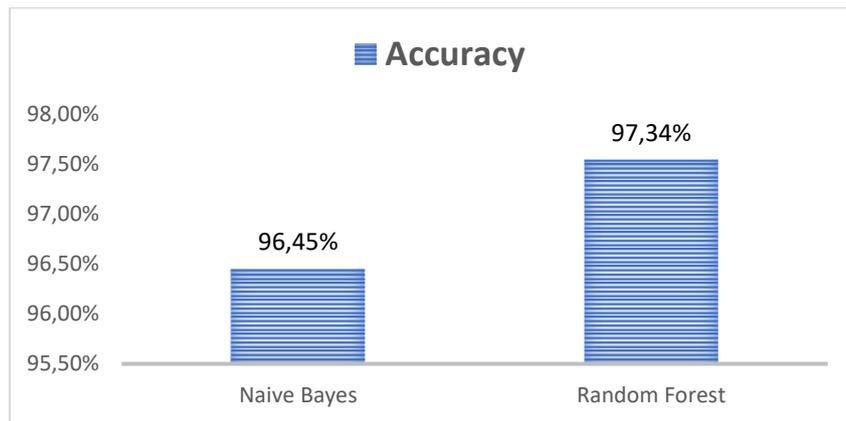
```

PerformanceVector:
accuracy: 97.34% +/- 0.42% (micro average: 97.34%)
ConfusionMatrix:
True:   1   0   2
1:     15277  389  28
0:      55   3145  0
2:      55    0   876
    
```

**Gambar 4.10** ConfusionMatrix Model dua

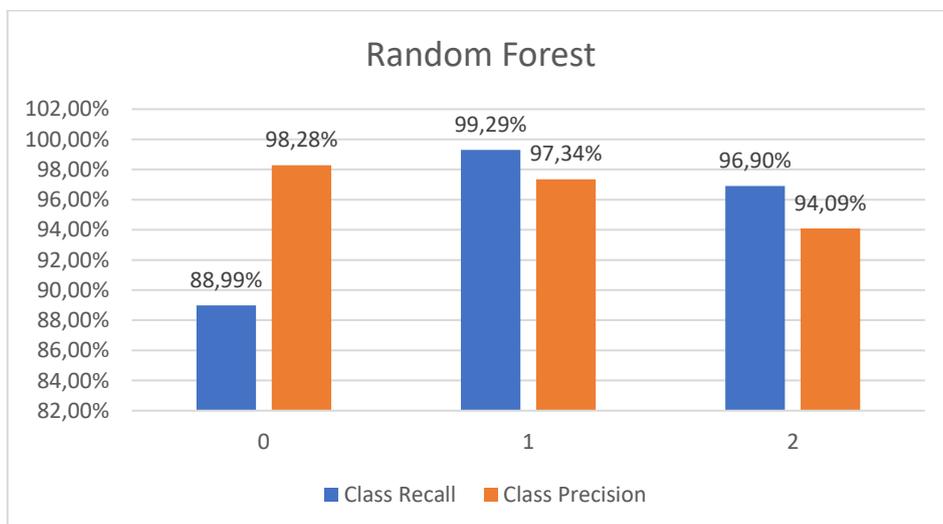
### 3.3. Pembahasan

Berdasarkan hasil pengujian model olgoritma diatas maka pada tahap ini dilakukan evaluasi dan pembahasan dari hasil pengolahan dataset produksi biskuit. Berdasarkan hasil pemrosesan yang sudah dilakukan, hasil prediksi menunjukkan bahwa untuk masing-masing variabel yang diprediksi nilai *Accuracy* dan *ConfusionMatrix* mengalami perbedaan di masing-masing model algoritma. Hal ini dikarenakan pada masing-masing varibel tersebut di uji dengan algoritma yang berbeda seperti yang sudah dicantumkan pada table di setiap model yang diujikan yang dimulai dari Model 1 sampai dengan Model 2. Sehingga pada saat pengujian menggunakan algoritma *Naïve Bayes* dan *Random Forest* dataset *General Report 2019-2023*, setiap algoritma memiliki nilai *Accuracy* dan *Confusion Matrix* dari masing-masing variabel yang diujikan memiliki hasil yang berbeda.



**Gambar 4.11** Grafik Perbandingan Algoritma pada *dataset* Produksi biskuit Tahun 2019-2023

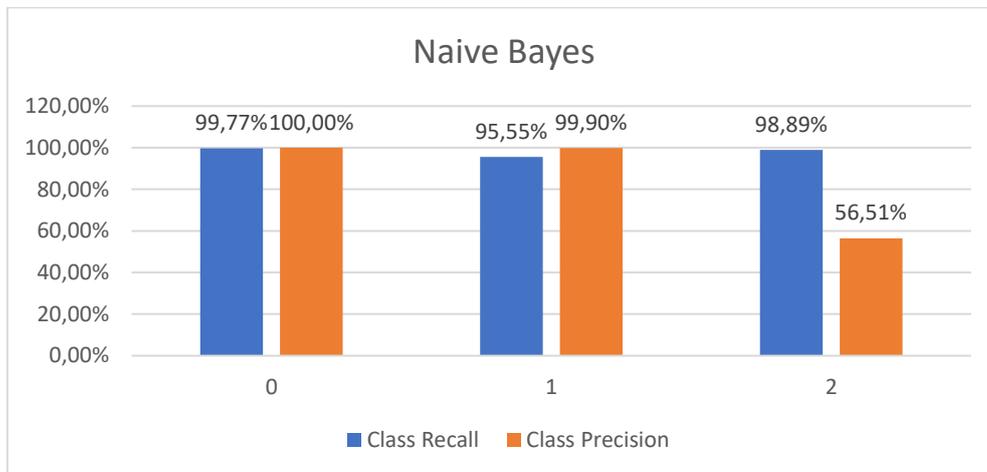
Pada gambar 4.19 dapat di lihat bahwa tingkat *Accuracy* dengan dataset *General Report 2019-2023* yang tertinggi adalah pada algoritma *Random Forest* yang mencapai **98.56%** dan *Naive Bayes* **96.45%** Sehingga dapat disimpulkan bahwa untuk *dataset General Report 2019-2023* algoritma dengan hasil *Accuray* terbaik adalah algoritma *Random Forest*.



**Gambar 4.12** Grafik Perbandingan *Recall* dan *Precision* pada Algoritma *Random Forest*

Pada gambar 4.20 dapat di lihat bahwa algoritma *Random forest* dengan dataset produksi biskuit tahun 2019-2023 yang tertinggi adalah pada Recall **99,29%** dan *Precisiion* **97,34%** pada *Class 1* diman *Class 1* adalah jumlah *Broken* dibawah 1000. Sehingga dapat disimpulkan bahwa algoritma *Random forest* untuk dataset produksi

biskuit tahun 2019-2023 *Recall* dan *Precision Class 1* dengan tingkat ketepatan prediksi yang tinggi.



**Gambar 4.13** Grafik Perbandingan *Recall* dan *Precision* pada Algoritma *Naïve Bayes*

Pada gambar 4.20 dapat di lihat bahwa algoritma *Naïve Bayes* dengan *dataset* produksi biskuit tahun 2019-2023 yang tertinggi adalah pada *Recall 99,77%* dan *Precision 100%* pada *Class 0* diman *Class 0* adalah jumlah *Broken 0*. Sehingga dapat disimpulkan bahwa algoritma *Naïve Bayes* untuk *dataset* produksi biskuit tahun 2019-2023 *Recall* dan *Precision Class 0* dengan tingkat ketepatan prediksi yang tinggi.

**Tabel 4.2** Perbandingan Hasil *Accuracy*

Model	Algoritma	Accuracy
Model 1	Naive Bayes	96,45%
Model 2	Random Forest	<b>97,34%</b>

Hasil prediksi dari lima model yang diujikan terlihat bahwa hasil dari *accuracy* algoritma ***Random Forest*** memiliki tingkat akurasi yang paling baik bila dibandingkan dengan algoritma *Naive Bayes* yang diujikan dalam penelitian ini.

#### 4. KESIMPULAN

Berdasarkan uraian-uraian yang telah dijabarkan pada bab sebelumnya, maka kesimpulan dari penelitian penulis yang berjudul “Analisis Data Produksi Biskuit Dengan Algoritma *Naive Bayes* Dan *Random Forest*” adalah :

1. Bahwa penelitian sebelumnya yang bertujuan untuk mengklasifikasi data menggunakan model yang berbeda-beda menunjukkan hasil dan *Accuracy* yang

berbeda antara satu dengan lainnya walaupun menggunakan variabel yang sama. Dengan adanya perbedaan hasil tersebut, perlu dilakukan perbandingan untuk mencari model dengan hasil akurasi tertinggi dan rata-rata error terendah.

2. Hasil penelitian menunjukkan bahwa algoritma *Random Forest* mempunyai hasil *Accuracy* tertinggi yaitu **97.34%** sedangkan *Naive Bayes* 96,45%.
3. Hasil klasifikasi jumlah broken produksi biskuit menggunakan data produksi biskuit tahun 2019-2023 menunjukkan bahwa jumlah broken produksi mengalami peningkatan ini dipengaruhi oleh faktor-faktor seperti suhu mesin, formulasi, bahan baku, kapasitas mesin dan lain-lain.

## 5. DAFTAR PUSTAKA

- [1] Agung Nugroho<sup>1</sup>, Yoga Religia<sup>2</sup>.(2021). Analisis Optimasi Algoritma Klasifikasi Naive Bayes menggunakan Genetic Algorithm dan Bagging, JURNAL RESTI (Rekayasa Sistem dan Teknologi Informasi), Vol. 5 No. 3 (2021) 504 - 510 ISSN Media Elektronik: 2580-0760
- [2] Alfian Zainal Macfud, Abdi Pandu Kusuma, Wahyu Dwi Puspitasari.(2023). Analisis Algoritma Naive Bayes Classifier (Nbc), Pada Klasifikasi Tingkat Minat Barang Di Toko Violet Cell, JATI (Jurnal Mahasiswa Teknik Informatika) Vol. 7 No. 1, Februari 2023
- [3] Angga Pebdika, Ruli Herdiana, Dodi Solihudin.(2023). Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima Pip, JATI (Jurnal Mahasiswa Teknik Informatika) Vol. 7 No. 1, Februari 2023
- [4] Chikitha Syahrika Arsyah<sup>1</sup>, Marina Elsera<sup>2</sup>.(2023).Implementasi Random Forest Dalam Melakukan Klasifikasi Kata Sarkasme Pada Media Sosial Facebook, Djtechno : Jurnal Teknologi Informasi, Vol. 4, No. 1 Juli 2023, E-ISSN: 2745-3758, P-ISSN : 2776-8546 DOI: 10.46576/djtechno
- [5] Dian Pramesti<sup>1</sup> & Wiga Maulana Baihaqi. (2022). Perbandingan Prediksi Jumlah Transaksi Ojek Online Menggunakan Regresi Linier dan Random Forest Generation Journal /Vol.7 No.3 (Special Issue) / e-ISSN: 2549-2233 / p-ISSN: 2580-4952.
- [6] Fajar Mu'Alim<sup>1</sup>), Rahmi Hidayati<sup>2</sup>).(2022).Implementasi Metode Random Forest Untuk Penjurusan Siswa Di Madrasah Aliyah Negeri Sintang, Jurnal JUPITER, Vol. 14 No. 1 Bulan April, Tahun 2022 , Hal. 116 – 125

- [7] Fandi Yulian Pamuji 1, Viry Puspaning Ramadhan 2.(2021).Komparasi Algoritma Random Forest Dan Decision Tree Untuk Memprediksi Keberhasilan Immunotherapy,Vol.7 No.1 Tahun 2021, pp 46-50, Jurnal Teknologi dan Manajemen Informatika <http://jurnal.unmer.ac.id/index.php/jtmi> P-ISSN: 1693-6604 E-ISSN: 2580-8044
- [8] Fauji Faisal Nugraha\*1, Iyan Sunandar2, Christina Juliane3.(2022). Penerapan Data Mining Dengan Metode Klasifikasi Menggunakan Algoritma C4.5, Jurnal Teknik Informatika dan Sistem Informasi ISSN 2407-4322, Vol. 9, No. 4, Desember 2022, Hal. 2862-2869 E- ISSN 2503-2933
- [9] Gabriela Militia Momole1, Evangs Mailoa2.(2022).Perbandingan Naïve Bayes Dan Random Forest Dalam Klasifikasi Bahasa Daerah, Jurnal Teknik Informatika dan Sistem Informasi ISSN 2407-4322, Vol. 9, No. 2, Juni 2022, Hal. 855-863 E- ISSN 2503-2933
- [10] Gde Agung Brahmana Suryanegara1, Adiwijaya2, Mahendra Dwifabri Purbolaksono3.(2021). Peningkatan Hasil Klasifikasi pada Algoritma Random Forest untuk Deteksi Pasien Penderita Diabetes Menggunakan Metode Normalisasi, JURNAL RESTI (Rekayasa Sistem dan Teknologi Informasi), Vol. 5 No. 1 (2021) 114 - 122 ISSN Media Elektronik: 2580-0760