

## BAB IV

### HASIL DAN PEMBAHASAN

Pada bab ini penulis akan membahas mengenai implementasi data yang digunakan. Data tersebut akan diimplementasikan menggunakan rapidminer menggunakan Adaboost dan Naïve Bayes.

#### 4.1. Persiapan Data

Data yang digunakan dalam penelitian ini adalah dataset dari kaggle <https://www.kaggle.com/datasets/andot03bsrc/dataset-predic-terkena-penyakit-paruparu/> dengan jumlah sebanyak 30.000 data yang menggunakan tools rapid miner. Atribut yang digunakan Usia, jenis kelamin, bekerja, rumah tangga, Aktivitas Begadang, Aktivitas Olahraga, Asuransi, Penyakit Bawaan. Metode yang digunakan dalam penelitian ini Adaboost dan Naïve Bayes.. Kumpulan data ini digunakan untuk memprediksi apakah seorang pasien kemungkinan besar menderita anemia. Data penyakit paru terdapat pada gambar 4.1 dibawah ini:

	▲ Aktivitas_Begada...	▲ Aktivitas_Olahraga	▲ Asuransi	▲ Penyakit_Bawaan
51%	Ya	Jarang	Ada	Ada
49%	Tidak	Sering	Tidak	Tidak
	Ya	Sering	Ada	Tidak
	Ya	Jarang	Ada	Ada
	Ya	Jarang	Ada	Tidak
	Tidak	Jarang	Ada	Ada
	Tidak	Sering	Tidak	Ada

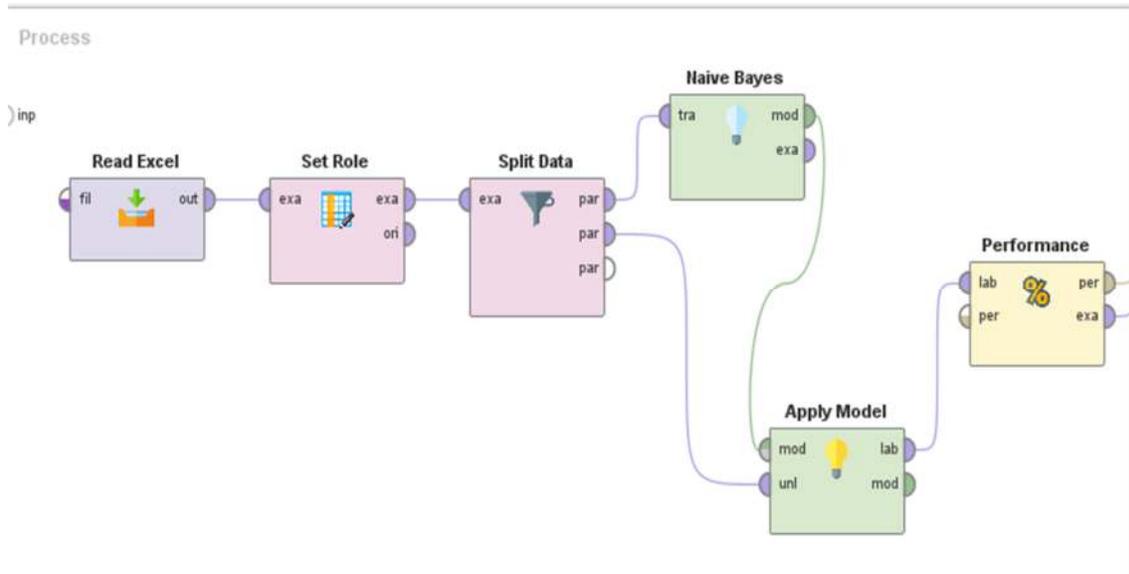
Gambar 4.1 Dataset Penyakit Paru-paru

#### 4.2. Pemodelan

Pemilihan dan penerapan teknik pemodelan yang sesuai dilakukan pada tahapan ini.. Pemodelan pada penelitian ini menggunakan teknik data mining Prediksi.

#### 4.2.1 Penelitian Menggunakan Algoritma Naïve Bayes

Penerapan data pada Rapidminer untuk Prediksi Penyakit Paru menggunakan algoritma Naïve Bayes ditunjukkan pada gambar dibawah ini:



Gambar 4.2 Penerapan Data Prediksi Paru Menggunakan Algoritma Naïve Bayes pada Rapidminer

Pada gambar 4.2 data yang sudah disiapkan diterapkan pada aplikasi rapidminer algoritma yang digunakan adalah Naïve Bayes Hasil eksperimen dapat kita lihat pada gambar 4.3 dibawah ini :

Table View  Plot View

accuracy: 87.22%

	true Ya	true Tidak	class precision
pred. Ya	5000	792	86.33%
pred. Tidak	741	5467	88.06%
class recall	87.09%	87.35%	

Table View  Plot View

recall: 87.35% (positive class: Tidak)

	true Ya	true Tidak	class precision
pred. Ya	5000	792	86.33%
pred. Tidak	741	5467	88.06%
class recall	87.09%	87.35%	

Table View  Plot View

precision: 88.06% (positive class: Tidak)

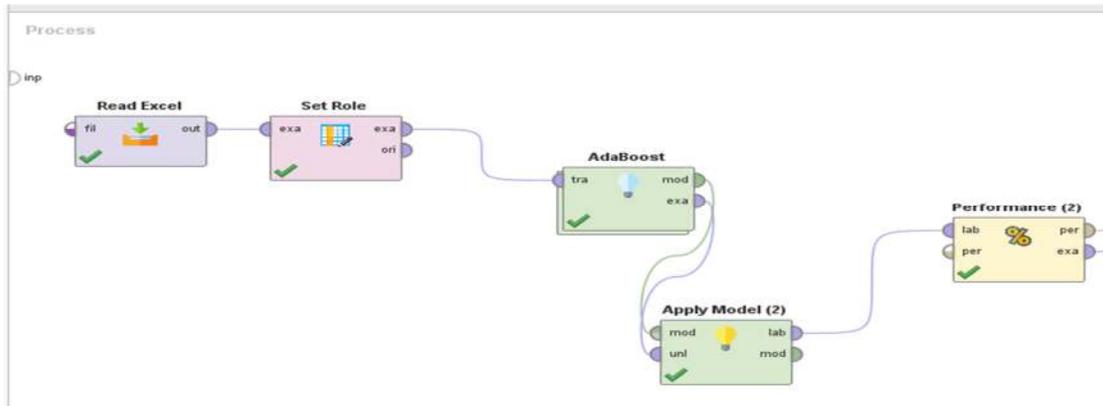
	true Ya	true Tidak	class precision
pred. Ya	5000	792	86.33%
pred. Tidak	741	5467	88.06%
class recall	87.09%	87.35%	

Gambar 4.3 Confusion Matrix Hasil Prediksi Paru Menggunakan  
Algoritma Naïve Bayes pada Rapidminer

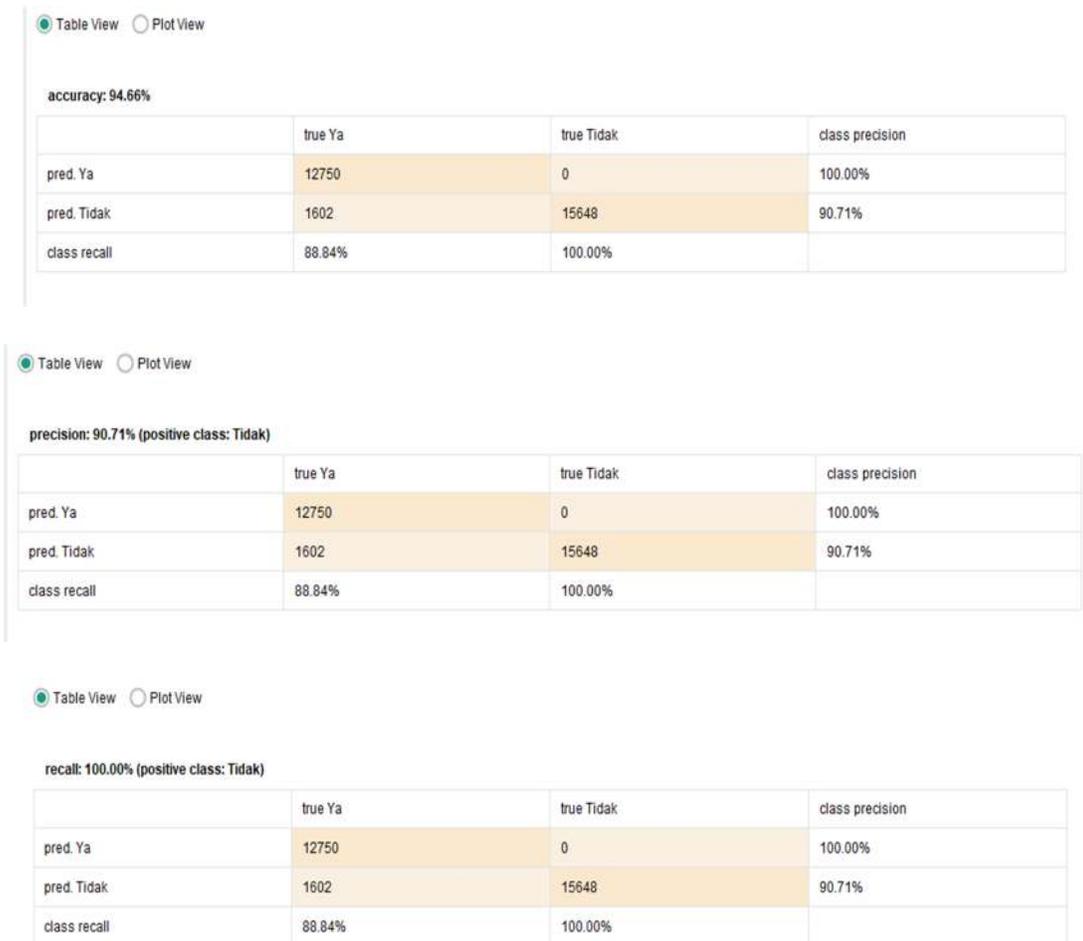
Gambar 4.3 adalah confusion matrix yang menunjukkan hasil eksperimen, didalam confusion matrix kita dapat melihat hasil akurasi, class presisi, dan class recall. Akurasi yang dihasilkan adalah 87,22 % class recall 87,35 % dan class presisi 88,06 %.

#### 4.2.1 Penelitian Menggunakan Adabost dan Algoritma Naïve Bayes

Metode yang digunakan menggunakan metode adabost dengan algoritma Naïve Bayes. Penerapan Metode dan algoritma pada rapidminer ditunjukkan pada gambar di bawah ini:



Gambar 4.4 Penerapan Data Prediksi Paru-Paru Menggunakan Metode Adabost dan Algoritma Naïve Bayes pada Rapidminer



Gambar 4.5 Confusion Matrix Hasil Prediksi Paru Menggunakan Algoritma Naïve Bayes dan Adabost pada Rapidminer

Gambar 4.5 adalah confusion matrix yang menunjukkan hasil eksperimen lanjutan, didalam confusion matrix kita dapat melihat hasil akurasi, class presisi, dan class recall. Akurasi yang dihasilkan adalah 94,66 % presisi 90,71%. dan recall 100,00 %.

$$Accuracy \frac{12750+15648}{12750+15648+1602+0} \times 100\% = \frac{28398}{30000} = 0,9466 \times 100\% = 94,66\%$$

$$Precision \frac{12750}{12750+1602} = \frac{12750}{14352} = 0,88837 = 88,84\%$$

$$Recall \frac{12750}{12750+0} = \frac{12750}{12750} = 100\%$$

Dari perhitungan diatas bisa diketahui bahwa Naïve Bayes merupakan algoritma klasifikasi probabilistik yang menggunakan teorema Bayes dengan asumsi independensi antar fitur. Ini sangat berguna untuk klasifikasi teks dan masalah lainnya. sedangkan AdaBoost (Adaptive Boosting) adalah metode ensemble yang menggabungkan beberapa model lemah (weak learners) menjadi satu model kuat (strong learner). Model lemah adalah model yang sedikit lebih baik dari tebakan acak, seperti pohon keputusan kecil.

Dalam eksperimen ini, penggunaan Adaboost bersama dengan Naive Bayes meningkatkan akurasi sebesar 7,44%. Ini berarti bahwa akurasi Naive Bayes awal adalah sekitar:  $94,66\% - 7,44\% = 87,22\%$ . Jadi, peningkatan akurasi dari 87,22% menjadi 94,66% menunjukkan kontribusi signifikan dari Adaboost dalam meningkatkan kinerja Naive Bayes

- Naive Bayes bekerja dengan menghitung probabilitas berdasarkan fitur individu dan kelas, dan sering digunakan sebagai model dasar dalam berbagai aplikasi.
- Adaboost meningkatkan kinerja model dengan menggabungkan model lemah dan memperbaiki kesalahan iterative.
- Kombinasi Naive Bayes dan Adaboost dapat meningkatkan akurasi model secara signifikan, seperti yang ditunjukkan oleh peningkatan 7,44% dalam eksperimen ini.

### 4.3.1. Perbandingan Penelitian Tidak Menggunakan Adabost dan Menggunakan Adabost

Perbandingan hasil eksperimen yang tidak menggunakan metode Adabost dan eksperimen yang menggunakan Adabost ditunjukkan pada tabel dibawah ini:

Tabel 4.3 Perbandingan Hasil Akurasi Penggunaan Adabost.

Penelitian	Tingkat Akurasi
Naïve Bayes	87,22 %
Naïve Bayes dan Adabost	94,66 %

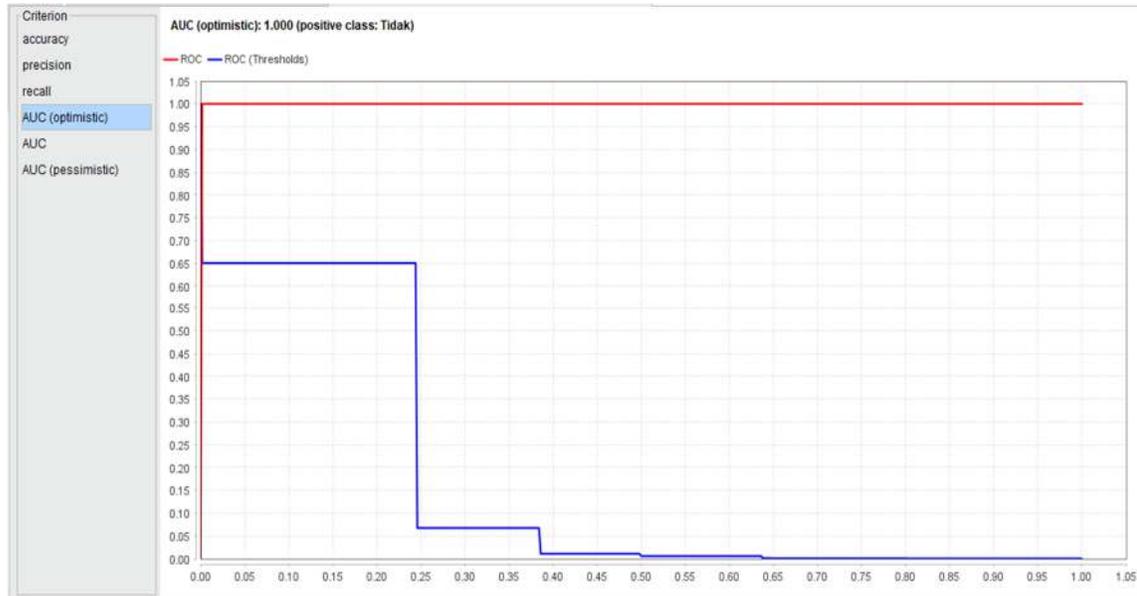
Hasil dari eksperimen ini untuk menghasilkan tingkat akurasi yang tinggi menggunakan naïve bayes dan adabost sebesar 94,66% dengan kenaikan akurasi sebesar 7,44%

Selain Confusion matrix untuk mengetahui kinerja dari eksperimen ini kita dapat mengandalkan kurva ROC-AUC yang dihasilkan. Perbandingan hasil Kurva ROC-AUC pada penelitian yang tidak menggunakan adabost dan yang menggunakan adabost dapat kita lihat pada gambar 4.6 dan 4.7 dibawah ini:



Gambar 4.6 Kurva ROC-AUC Tidak Menggunakan Adabost

Gambar 4.6 menunjukkan kinerja algoritma dalam prediksi paru, kinerja algoritma pada eksperimen ini sangat baik dengan menghasilkan Area Under Curva (AUC) 85%.



Gambar 4.7 Kurva ROC-AUC Menggunakan Adabost

Gambar 4.6 menunjukkan kinerja algoritma dalam prediksi paru, kinerja algoritma menggunakan adabost pada eksperimen ini sangat baik dengan menghasilkan Area Under Curva (AUC) 100%.

### 4.3 Evaluasi

Pemodelan yang telah dilakukan dengan melakukan eksperimen menggunakan algoritma naïve bayes kurang menghasilkan akurasi yang baik. Dan setelah penggunaan adabost berhasil meningkatkan akurasi dalam prediksi penyakit paru-paru. Selain itu kinerja algoritma pada penelitian ini juga menunjukkan performa yang sangat baik sekali.