

Manual Program

Perbandingan Metode Naive Bayes, GBM
dan KNN Terhadap Pasien Penyakit ISPA
Studi Kasus: Rumah Sakit Royal Taruma



UNTAR

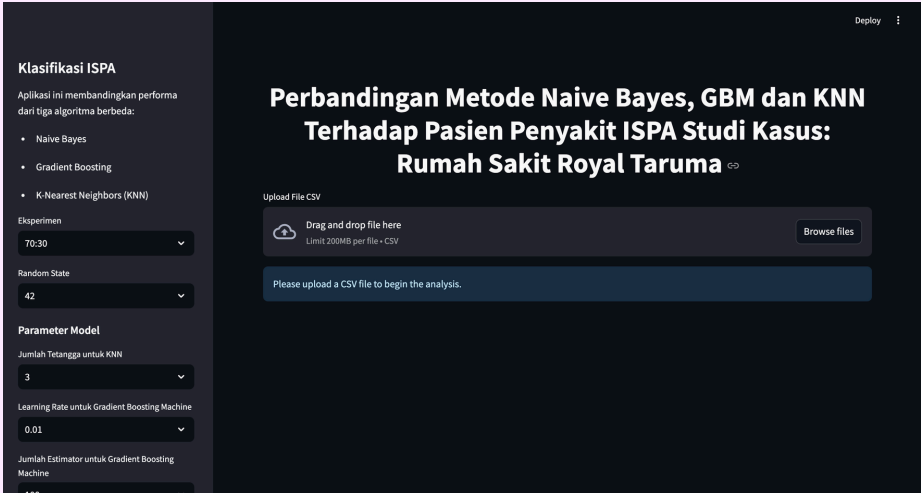
Universitas Tarumanagara

Apriyanto Chandra

535210032

Antarmuka Pengguna

1. Halaman Home



The screenshot displays the user interface of a web application. On the left is a dark sidebar with the title "Klasifikasi ISPA". Below the title, it states "Aplikasi ini membandingkan performa dari tiga algoritma berbeda:" followed by a list of algorithms: Naive Bayes, Gradient Boosting, and K-Nearest Neighbors (KNN). Further down, there are sections for "Eksperimen" with a "70:30" dropdown, "Random State" with a "42" dropdown, and "Parameter Model" which includes "Jumlah Tetangga untuk KNN" (set to 3), "Learning Rate untuk Gradient Boosting Machine" (set to 0.01), and "Jumlah Estimator untuk Gradient Boosting Machine" (set to 100). The main content area has a title "Perbandingan Metode Naive Bayes, GBM dan KNN Terhadap Pasien Penyakit ISPA Studi Kasus: Rumah Sakit Royal Taruma" with a link icon. Below the title is an "Upload File CSV" section with a drag-and-drop area that says "Drag and drop file here" and "Limit 200MB per file • CSV", a "Browse files" button, and a message "Please upload a CSV file to begin the analysis."

Di halaman home, terdapat area untuk mengunggah file data. *User* bisa langsung menarik (drag) dan melepaskan (drop) file CSV ke dalam area tersebut atau memilih file secara manual dengan menekan tombol "Browse files".

Pastikan file CSV berukuran tidak lebih dari 200MB dan memiliki kolom-kolom data yang sesuai dengan kebutuhan aplikasi untuk memastikan analisis dapat berjalan dengan baik.

2. Sidebar

Klasifikasi ISPA

Aplikasi ini membandingkan performa dari tiga algoritma berbeda:

- Naive Bayes
- Gradient Boosting
- K-Nearest Neighbors (KNN)

Eksperimen

70:30

Random State

42

Parameter Model

Jumlah Tetangga untuk KNN

3

Learning Rate untuk Gradient Boosting Machine

0.01

Jumlah Estimator untuk Gradient Boosting Machine

100

Maksimal Depth untuk Decision Tree

3

Di bagian kiri layar, *user* akan melihat beberapa pilihan pengaturan untuk eksperimen klasifikasi disini, *user* bisa memilih rasio data latih dan data uji, misalnya 70:30, yang berarti 70% data akan digunakan untuk pelatihan model dan 30% untuk pengujian. Random State: Pilih angka random state, seperti angka 42, untuk menjaga agar hasil eksperimen tetap konsisten setiap kali dilakukan. *User* dapat mengatur beberapa parameter model agar eksperimen lebih sesuai dengan kebutuhan pada algoritma k-nearest neighbors dan gradient boosting machine

3. Submit File CSV

Klasifikasi ISPA

Aplikasi ini membandingkan performa dari tiga algoritma berbeda:

- Naive Bayes
- Gradient Boosting
- K-Nearest Neighbors (KNN)

Eksperimen

70:30

Random State

42

Parameter Model

Jumlah Tetangga untuk KNN

3

Learning Rate untuk Gradient Boosting Machine

0.01

Jumlah Estimator untuk Gradient Boosting Machine

100

Deploy

Perbandingan Metode Naive Bayes, GBM dan KNN Terhadap Pasien Penyakit ISPA Studi Kasus: Rumah Sakit Royal Taruma

Upload File CSV

Drag and drop file here
Limit 200MB per file • CSV

Browse files

testing.csv 28.6KB

Submit

Please upload a CSV file to begin the analysis.

Setelah semua pengaturan sudah selesai dan file data CSV berhasil diunggah, web akan memproses data dan melakukan analisis klasifikasi menggunakan ketiga algoritma. Hasil performa dari Naive Bayes, Gradient Boosting, dan KNN akan ditampilkan untuk mempermudah Anda dalam membandingkan masing-masing model.

4. Hasil Analisis

Klasifikasi ISPA

Aplikasi ini membandingkan performa dari tiga algoritma berbeda:

- Naive Bayes
- Gradient Boosting
- K-Nearest Neighbors (KNN)

Ekspirimen

70:30

Random State

42

Parameter Model

Jumlah Tetangga untuk KNN

3

Learning Rate untuk Gradient Boosting Machine

0.01

Jumlah Estimator untuk Gradient Boosting Machine

Deploy

Upload File CSV

Drag and drop file here

Limit 200MB per file • CSV

Browse files

testing.csv

28.6KB

Submit

Data Mentah

	id	age	gender	sesak_napas	suhu_tubuh	hr	musal	rr	Spo2	batuk	pilek	demam	ispa
0	1	63	lk	iya	36.2	93 bpm	tidak	22	97%	tidak	tidak	tidak	tidak
1	2	35	pr	tidak	36.1	96 bpm	tidak	20	98%	iya	iya	tidak	tidak
2	3	6	lk	iya	36.1	120 bpm	iya	22	98%	iya	iya	iya	iya
3	4	3	lk	iya	36	103 bpm	tidak	22	98%	iya	iya	tidak	iya
4	5	1	pr	iya	36.5	120 bpm	tidak	22	98%	iya	iya	iya	iya
5	6	38	lk	tidak	36	92 bpm	tidak	20	97%	iya	iya	tidak	tidak
6	7	14	pr	tidak	36.3	93 bpm	tidak	20	97%	iya	tidak	tidak	tidak
7	8	21	lk	tidak	36.3	73 bpm	tidak	20	100%	tidak	tidak	iya	tidak
8	9	38	pr	tidak	36.4	103 bpm	tidak	20	100%	tidak	iya	iya	iya
9	10	61	lk	tidak	36.4	89 bpm	iya	20	97%	iya	iya	iya	iya

Pada bagian ini, data mentah yang diunggah oleh pengguna akan ditampilkan dalam bentuk tabel. Tabel ini memungkinkan pengguna untuk meninjau dan memahami struktur serta karakteristik data sebelum proses preprocessing dilakukan. Setiap kolom dalam tabel ini mewakili fitur-fitur yang ada dalam dataset.

5. Preprocessing

Klasifikasi ISPA

Aplikasi ini membandingkan performa dari tiga algoritma berbeda:

- Naive Bayes
- Gradient Boosting
- K-Nearest Neighbors (KNN)

Eksperimen

70:30

Random State

42

Parameter Model

Jumlah Tetangga untuk KNN

3

Learning Rate untuk Gradient Boosting Machine

0.01

Jumlah Estimator untuk Gradient Boosting Machine

Preprocessing Data

Missing Values sebelum preprocessing:

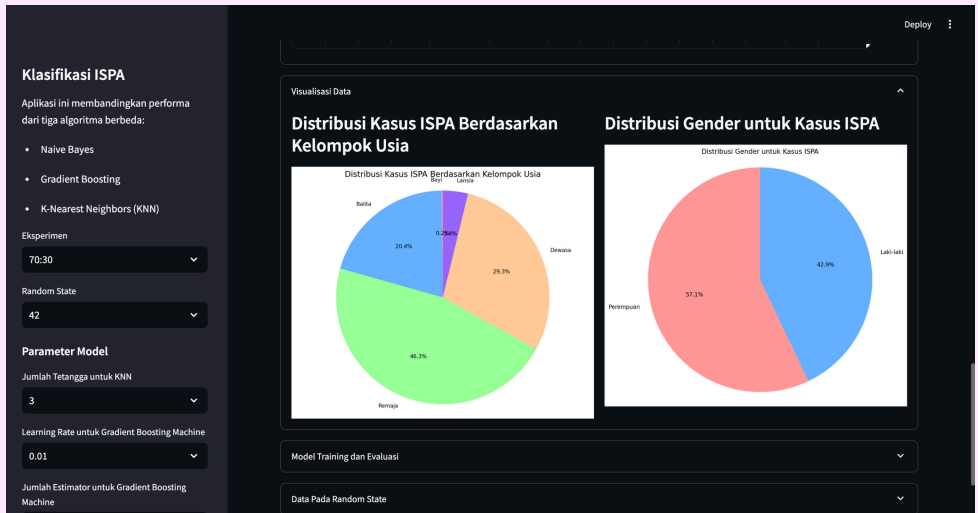
	0
id	0
age	0
gender	0
sesak_napas	122
suhu_tubuh	20
hr	58
mual	45
rr	124
Spo2	87
batuk	0

Missing Values setelah preprocessing:

	0
id	0
age	0
gender	0
sesak_napas	0

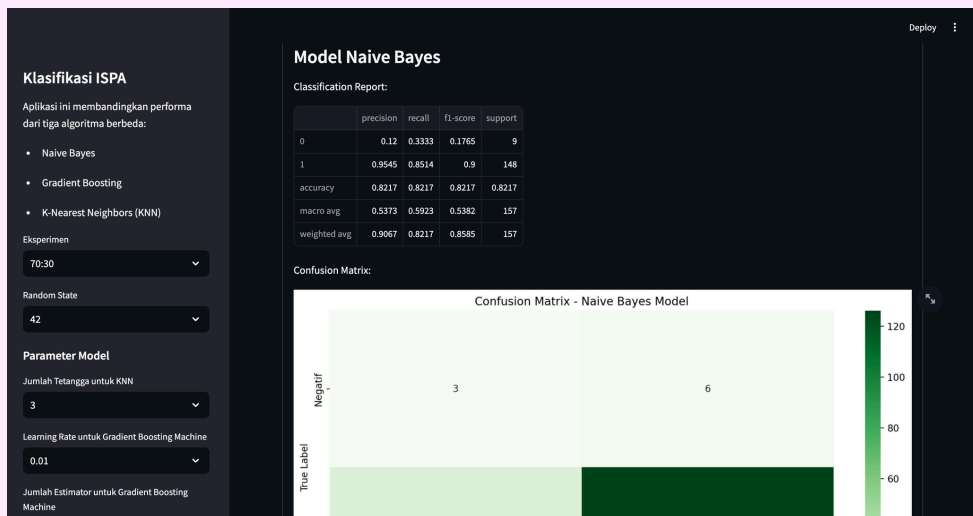
Setelah proses preprocessing selesai, data yang telah diproses akan ditampilkan di tabel "Data setelah preprocessing". Pada tahap ini, berbagai transformasi telah dilakukan untuk mempersiapkan data agar sesuai dengan kebutuhan model. Data yang telah dipreproses ini sudah dalam format yang sesuai untuk diolah oleh model Naive Bayes, Gradient Boosting, dan KNN, yang akan digunakan untuk melakukan klasifikasi pada penyakit ISPA.

6. Visualisasi Data



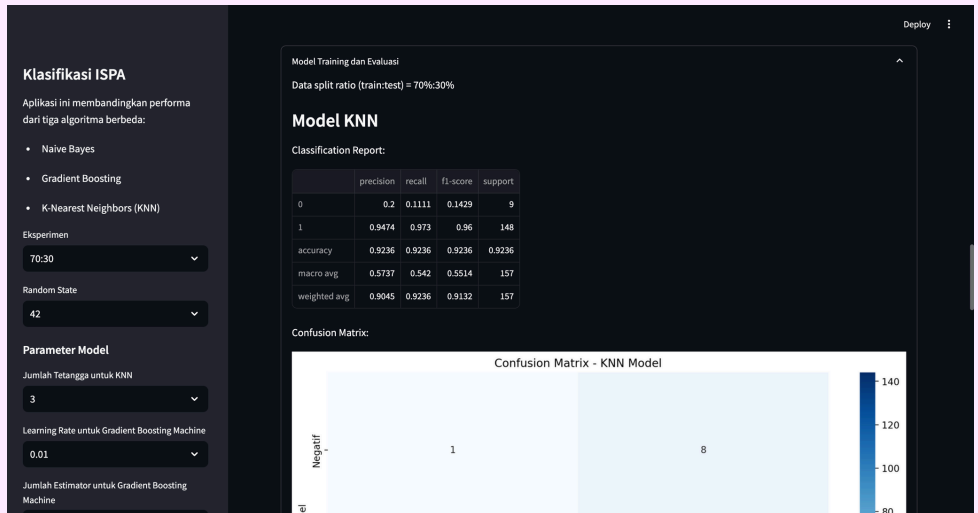
User dapat mengklik bagian "Visualisasi Data" untuk melihat berbagai grafik yang menggambarkan distribusi kasus ISPA berdasarkan kelompok usia, dan distribusi gender untuk kasus ISPA. Fitur ini menyediakan tampilan visual yang penting untuk memahami data.

7. Model Naive Bayes



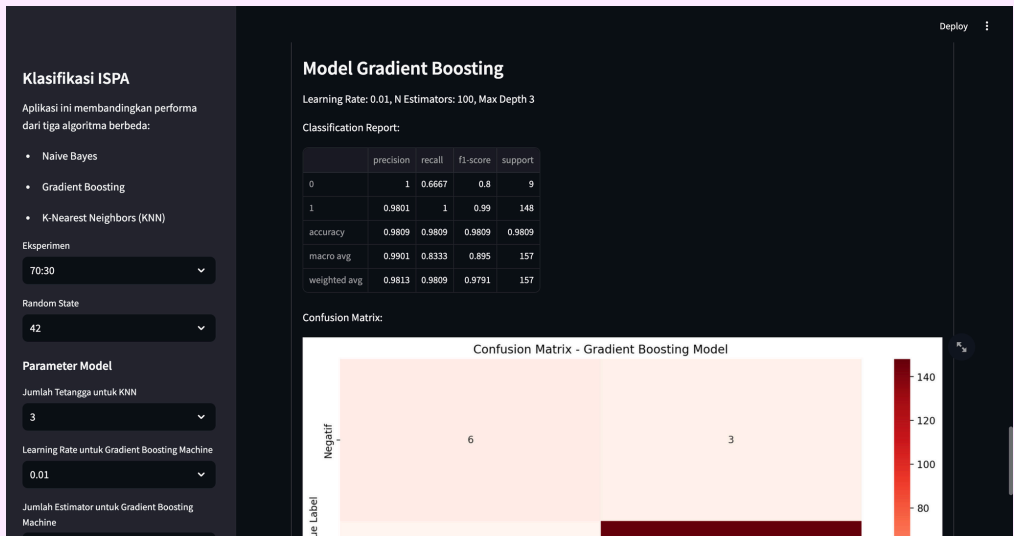
Pada bagian ini, model Naive Bayes dilatih menggunakan data yang telah melalui tahap preprocessing untuk memastikan bahwa model dapat bekerja secara optimal dengan data yang bersih dan konsisten. Setelah proses pelatihan selesai, performa model akan dievaluasi menggunakan metrik-metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Selain itu, *confusion matrix* juga akan ditampilkan untuk memberikan gambaran visual mengenai hasil prediksi model. Dengan melihat matriks ini, pengguna dapat memahami tingkat kesalahan prediksi yang terjadi, serta apakah model cenderung melakukan kesalahan pada kelas tertentu.

8. Model K-Nearest Neighbors



Pada bagian ini, model k-Nearest Neighbors dilatih menggunakan data yang telah melalui tahap preprocessing untuk memastikan bahwa model dapat bekerja secara optimal dengan data yang bersih dan konsisten. Setelah proses pelatihan selesai, performa model akan dievaluasi menggunakan metrik-metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Selain itu, *confusion matrix* juga akan ditampilkan untuk memberikan gambaran visual mengenai hasil prediksi model. Dengan melihat matriks ini, pengguna dapat memahami tingkat kesalahan prediksi yang terjadi, serta apakah model cenderung melakukan kesalahan pada kelas tertentu.

9. Model Gradient Boosting Machine



Pada bagian ini, model Gradient Boosting Machine dilatih menggunakan data yang telah melalui tahap preprocessing untuk memastikan bahwa model dapat bekerja secara optimal dengan data yang bersih dan konsisten. Setelah proses pelatihan selesai, performa model akan dievaluasi menggunakan metrik-metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Selain itu, *confusion matrix* juga akan ditampilkan untuk memberikan gambaran visual mengenai hasil prediksi model. Dengan melihat matriks ini, pengguna dapat memahami tingkat kesalahan prediksi yang terjadi, serta apakah model cenderung melakukan kesalahan pada kelas tertentu.

10. Random State

Klasifikasi ISPA

Aplikasi ini membandingkan performa dari tiga algoritma berbeda:

- Naive Bayes
- Gradient Boosting
- K-Nearest Neighbors (KNN)

Eksperimen

70:30

Random State

42

Parameter Model

Jumlah Tetangga untuk KNN

3

Learning Rate untuk Gradient Boosting Machine

0.01

Jumlah Estimator untuk Gradient Boosting Machine

100

Deploy

Data Pada Random State

Data Latih (X_train):

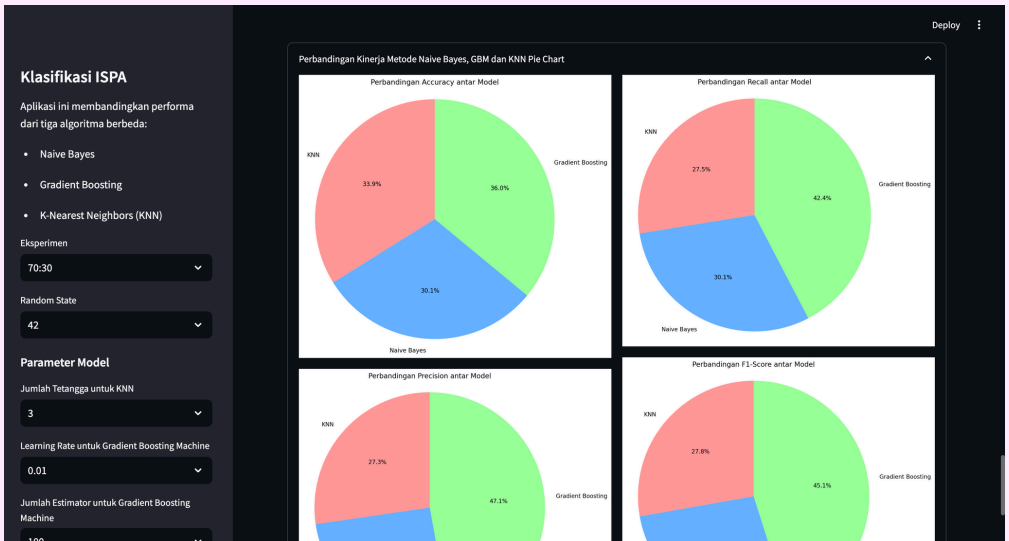
	age	gender	sesak_napas	suhu_tubuh	hr	musl	rr	Spo2	batuk	pilek	demam	age_group
231	12	0	1	36.7	149	1	22	97	1	1	1	2
370	8	1	1	36.5	110	0	21	99	1	1	1	2
94	33	0	0	36.6	76	0	20	100	1	1	1	3
209	1	0	1	36	129	0	30	98	1	0	0	1
497	74	1	0.5567	36	83	0	22.4332	97	1	1	0	4
23	12	0	0.5567	37.181	120	1	22.4332	97.7903	1	1	1	2
322	25	1	0	36	88	0	20	100	1	0	1	3
271	32	1	0	36.4	91	1	20	96	0	1	1	3
425	16	1	1	36.4	76	1	22	99	1	1	0	2
5	38	1	0	36	92	0	20	97	1	1	0	3

Label Latih (y_train):

	ispa
231	1
370	1
94	1
209	1

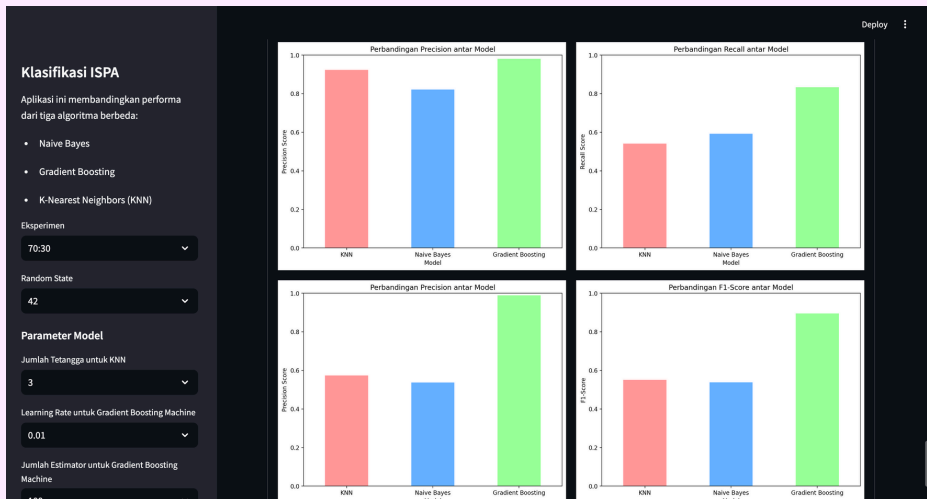
Pada bagian ini, aplikasi ini menyediakan tampilan yang interaktif untuk memvisualisasikan data latih (X_train), label (y_train) dan data uji (X_test), label (y_test) yang dihasilkan berdasarkan pengaturan random_state.

11. Evaluasi Matrix Menggunakan Pie Chart



Bagian ini menampilkan perbandingan kinerja ketiga metode klasifikasi, yaitu Naive Bayes, Gradient Boosting, dan K-Nearest Neighbors (KNN), dalam bentuk diagram pie berdasarkan metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Diagram pie ini memberikan gambaran visual mengenai persentase kinerja masing-masing model, memungkinkan user untuk secara intuitif memahami seberapa baik setiap model bekerja dalam mengklasifikasikan data ISPA.

12. Evaluasi Matrix Menggunakan Bar Chart



Bagian ini menampilkan perbandingan kinerja ketiga metode klasifikasi, yaitu Naive Bayes, Gradient Boosting, dan K-Nearest Neighbors (KNN), dalam bentuk diagram batang berdasarkan metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Diagram batang ini memberikan gambaran visual mengenai persentase kinerja masing-masing model, memungkinkan *user* untuk secara intuitif memahami seberapa baik setiap model bekerja dalam mengklasifikasikan data ISPA.

Catatan

This image shows a single sheet of white paper with horizontal black ruling lines. The paper has rounded corners at the top and bottom. There are 18 horizontal lines in total, spaced evenly apart. The lines extend across the width of the page, leaving small margins on the left and right sides. The background behind the paper is a light purple color.



UNTAR
Universitas Tarumanagara