

PERBANDINGAN KINERJA METODE *NAÏVE BAYES* DAN *RANDOM FOREST* UNTUK KLASIFIKASI RISIKO SAHAM PERBANKAN

NAUFAL IRSYAD AS-SAMI



**PROGRAM STUDI AKTUARIA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Perbandingan Kinerja Metode *Naïve Bayes* dan *Random Forest* untuk Klasifikasi Risiko Saham Perbankan” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juli 2025

Naufal Irsyad As-Sami
NIM. G5402211068



ABSTRAK

NAUFAL IRSYAD AS-SAMI. Perbandingan Kinerja Metode *Naïve Bayes* dan *Random Forest* untuk Klasifikasi Risiko Saham Perbankan. Dibimbing oleh I WAYAN MANGKU dan RUHIYAT.

Ketidakpastian risiko pada investasi saham menuntut pendekatan yang andal dalam menentukan keputusan investasi. Salah satu pendekatan yang dapat digunakan adalah metode klasifikasi seperti *naïve Bayes* dan *random forest*. Penelitian ini bertujuan membandingkan kinerja metode *naïve Bayes* dan *random forest* dalam mengklasifikasikan kelas risiko berupa risiko tinggi dan risiko rendah pada 25 saham sektor perbankan di Indonesia. Namun, tidak ada kriteria yang pasti dalam mengelompokkan kelas risiko saham. Dengan demikian, pengelompokan kelas data aktual memanfaatkan volatilitas sebagai gambaran fluktuatif dari risiko dengan nilai mediannya sebagai pembatas untuk kelas risiko tinggi dan kelas risiko rendah. Kedua metode klasifikasi yang digunakan bekerja berdasarkan variabel prediktor berupa *return* dan *price range*. *Naïve Bayes* memodelkan distribusi *Gaussian* untuk menghitung probabilitas kelas, sedangkan *random forest* membangun serangkaian pohon keputusan untuk menentukan kelas klasifikasi berdasarkan pemungutan suara mayoritas. Evaluasi kinerja kedua metode klasifikasi terhadap data uji menggunakan metrik akurasi, presisi, *recall*, dan *f1-score*. Hasil yang diperoleh menunjukkan bahwa *random forest* dengan akurasi tahunan sebesar 0.8433, presisi 0.7769, *recall* 0.8246, dan *f1-score* 0.8000 adalah metode klasifikasi yang lebih baik secara keseluruhan dibandingkan *naïve Bayes* dengan akurasi tahunan sebesar 0.7833, presisi 0.8182, *recall* 0.5526, dan *f1-score* 0.6597.

Kata kunci: kelas risiko, klasifikasi, *naïve Bayes*, *random forest*.

ABSTRACT

NAUFAL IRSYAD AS-SAMI. Comparison of the Performance of Naïve Bayes and Random Forest Methods for Banking Stock Risk Classification. Supervised by I WAYAN MANGKU and RUHIYAT.

The uncertainty of risk in stock investment demands a reliable approach to support investment decision-making. One such approach is the use of classification methods, such as naïve Bayes and random forest. This study aims to compare the performance of the naïve Bayes and random forest methods in classifying risk classes, namely high risk and low risk, for 25 banking sector stocks in Indonesia. However, there are no definite criteria for grouping stock risk classes. Therefore, the actual data classes are grouped using volatility as a representation of risk fluctuations, with its median value serving as the threshold between high-risk and low-risk classes. Both classification methods work based on predictor variables, namely return and price range. Naïve Bayes models a Gaussian distribution to calculate class probabilities, while random forest builds an ensemble of decision trees to determine the classification outcome based on majority voting. The performance of both classification methods on the test data is evaluated using accuracy, precision, recall, and f1-score metrics. The results show that random forest, with an annual accuracy of 0.8433, precision of 0.7769, recall of 0.8246, and f1-score of 0.8000, is an overall better classification method compared to naïve Bayes, which achieves an annual accuracy of 0.7833, precision of 0.8182, recall of 0.5526, and f1-score of 0.6597.

Keywords: classification, naïve Bayes, random forest, risk classes.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2025
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

PERBANDINGAN KINERJA METODE *NAÏVE BAYES* DAN *RANDOM FOREST* UNTUK KLASIFIKASI RISIKO SAHAM PERBANKAN

NAUFAL IRSYAD AS-SAMI

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Aktuaria

**PROGRAM STUDI AKTUARIA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:
Ir. N. K. Kutha Ardana, M.Sc.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Judul Skripsi : Perbandingan Kinerja Metode *Naïve Bayes* dan *Random Forest*
untuk Klasifikasi Risiko Saham Perbankan

Nama : Naufal Irsyad As-Sami
NIM : G5402211068

@Hak cipta milik IPB University

Disetujui oleh

Pembimbing 1:
Prof. Dr. Ir. I Wayan Mangku, M.Sc.

Pembimbing 2:
Ruhayat, S.Si., M.Si., M.Act.Sc.

Diketahui oleh

Ketua Program Studi:
Dr. Ir. I Gusti Putu Purnaba, DEA.
NIP. 196512181990021001

Tanggal Ujian: 9 Juli 2025

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah *Subhanaahu Wa Ta'ala* atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan April 2025 sampai bulan Juni 2025 ini ialah “Perbandingan Kinerja Metode *Naïve Bayes* dan *Random Forest* untuk Klasifikasi Risiko Saham Perbankan”. Ucapan terima kasih penulis sampaikan kepada:

1. Kedua orang tua, Bapak Asep Hikayat dan Ibu Yanti Sri Gustianti, serta ketiga adik penulis, Nufail Raimy Aziz, Nawaf Ruzain Hakam, dan Najwan Fariz Hafidz, yang tidak berhenti memberikan dukungan penuh dan kasih sayang selama menempuh pendidikan dan penulisan karya ilmiah.
2. Dosen pembimbing pertama, Bapak Prof. Dr. Ir. I Wayan Mangku M.Sc., dan dosen pembimbing kedua, Bapak Ruhiyat, S.Si., M.Si., M.Act.Sc., atas segala arahan, saran, motivasi, dan ilmu pengetahuan yang diberikan kepada penulis selama menyelesaikan karya ilmiah ini.
3. Dosen penguji, Bapak Ir. N. K. Kutha Ardana, M.Sc., yang memberikan kritik dan saran selama pelaksanaan Seminar dan Sidang Tugas Akhir.
4. Seluruh dosen pengajar dan tenaga kependidikan Program Studi Aktuaria SSMI IPB atas segala ilmu dan bantuan yang telah diberikan selama perkuliahan dan penulisan karya ilmiah ini.
5. Teman-teman gimik seperjuangan saya yaitu Hadi, Arif, Ravy, Haidar, Ikiw, Dita, Nabilah, Dhea, dan Kanaya yang selalu ada sebagai tempat berbagi cerita dan memberi dukungan selama masa studi ini.
6. Teman-teman Aktuaria 58, ASSA IPB, 2 SAM Kesayangan, SAMORSE, Dunia Supporteran, KKNT Mesuji 02, Bocil Actuary PSLA yang telah memberikan pengalaman berharga dan dukungan selama masa studi ini.
7. Seluruh pihak yang tidak dapat penulis sebutkan satu per satu, yang telah membantu saya dalam proses penyusunan karya ilmiah ini.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Juli 2025

Naufal Irsyad As-Sami



Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

DAFTAR ISI

DAFTAR TABEL	xi
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xi
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	1
1.3 Tujuan	1
TINJAUAN PUSTAKA	2
2.1 Investasi	2
2.2 Saham	2
2.3 <i>Return</i> Saham	2
2.4 <i>Price Range</i>	3
2.5 Volatilitas Harga Saham	3
2.6 Kelas Risiko Saham	3
2.7 <i>Naïve Bayes</i>	3
2.8 <i>Random Forest</i>	4
2.9 Evaluasi Kinerja Model	5
III METODE	7
3.1 Sumber Data	7
3.2 Perangkat Lunak	7
3.3 Prosedur Analisis Data	7
IV HASIL DAN PEMBAHASAN	9
4.1 Eksplorasi Data	9
4.2 Pembagian Data	10
4.3 <i>Naïve Bayes</i>	13
4.4 <i>Random Forest</i>	16
4.5 Evaluasi Kinerja Kedua Model	20
V SIMPULAN DAN SARAN	27
5.1 Simpulan	27
5.2 Saran	27
DAFTAR PUSTAKA	28
LAMPIRAN	30
RIWAYAT HIDUP	41

DAFTAR TABEL

1	<i>Confusion matrix</i>	5
2	Pembagian data latih dan data uji	10
3	<i>Confusion matrix naïve Bayes</i> pada kuartal I tahun 2024	20
4	<i>Confusion matrix naïve Bayes</i> pada kuartal II tahun 2024	20
5	<i>Confusion matrix naïve Bayes</i> pada kuartal III tahun 2024	21
6	<i>Confusion matrix naïve Bayes</i> pada kuartal IV tahun 2024	21
7	<i>Confusion matrix random forest</i> pada kuartal I tahun 2024	22
8	<i>Confusion matrix random forest</i> pada kuartal II tahun 2024	22
9	<i>Confusion matrix random forest</i> pada kuartal III tahun 2024	23
10	<i>Confusion matrix random forest</i> pada kuartal IV tahun 2024	23
11	Kinerja model <i>naïve Bayes</i> pada data uji bulanan tahun 2024	24
12	Kinerja model <i>random forest</i> pada data uji bulanan tahun 2024	24
13	<i>Confusion matrix naïve Bayes</i> dan <i>random forest</i> pada data uji tahunan	25
14	Kinerja model <i>naïve Bayes</i> dan <i>random forest</i> pada data uji tahunan	26

DAFTAR GAMBAR

1	Diagram alir penelitian	8
2	<i>Heatmap</i> korelasi antar variabel prediktor	9
3	Plot sebaran data uji sebelum pemodelan pada kuartal I tahun 2024	11
4	Plot sebaran data uji sebelum pemodelan pada kuartal II tahun 2024	11
5	Plot sebaran data uji sebelum pemodelan pada kuartal III tahun 2024	12
6	Plot sebaran data uji sebelum pemodelan pada kuartal IV tahun 2024	13
7	Plot sebaran data uji model <i>naïve Bayes</i> pada kuartal I tahun 2024	14
8	Plot sebaran data uji model <i>naïve Bayes</i> pada kuartal II tahun 2024	14
9	Plot sebaran data uji model <i>naïve Bayes</i> pada kuartal III tahun 2024	15
10	Plot sebaran data uji model <i>naïve Bayes</i> pada kuartal IV tahun 2024	16
11	Plot sebaran data uji model <i>random forest</i> pada kuartal I tahun 2024	17
12	Plot sebaran data uji model <i>random forest</i> pada kuartal II tahun 2024	18
13	Plot sebaran data uji model <i>random forest</i> pada kuartal III tahun 2024	19
14	Plot sebaran data uji model <i>random forest</i> pada kuartal IV tahun 2024	19

DAFTAR LAMPIRAN

1	Daftar 25 saham yang digunakan dalam penelitian	31
2	Kode R untuk klasifikasi <i>naïve Bayes</i>	32
3	Kode R untuk klasifikasi <i>random forest</i>	36
4	Perbandingan kelas risiko sebelum dan setelah klasifikasi dengan kedua metode	40