



**PERBANDINGAN *SUPPORT VECTOR MACHINE*, *RANDOM FOREST* DAN
NAÏVE BAYES PADA KLASIFIKASI SUARA BATUK COVID-19**

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Sarjana Strata-1 Ilmu Komputer**

**Oleh
BRIGITHA PRICILIA RAHMI
NIM. 1611016220005**

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
JUNI 2023**



**PERBANDINGAN *SUPPORT VECTOR MACHINE*, *RANDOM FOREST* DAN
NAÏVE BAYES PADA KLASIFIKASI SUARA BATUK COVID-19**

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Sarjana Strata-1 Ilmu Komputer
Oleh**

BRIGITHA PRICILIA RAHMI

NIM. 1611016220005

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

JUNI 2023

SKRIPSI

**PERBANDINGAN *SUPPORT VECTOR MACHINE*, *RANDOM FOREST* DAN
NAÏVE BAYES PADA KLASIFIKASI SUARA BATUK COVID-19**

Oleh:

BRIGITHA PRICILIA RAHMI

NIM. 1611016220005

Telah dipertahankan di depan Dosen Penguji pada tanggal 22 Juni 2023.

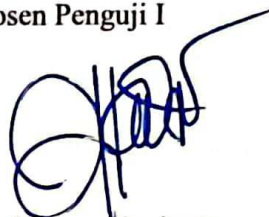
Susunan Dosen Penguji:

Pembimbing I



**Mohammad Reza Faisal, S.T.,
M.T., Ph.D.**
NIP. 197612202008121001

Dosen Penguji I



**Dodon Turianto Nugrahadi,
S.Kom., M.Eng.**
NIP. 198001122009121002

Pembimbing II



Dwi Kartini, S.Kom., M.Kom.
NIP. 198704212012122003

Dosen Penguji II



Muhammad Itqan Mazdadi, S.Kom., M.Kom.
NIP. 199006122019031013



PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diberikan orang lain, kecuali yang secara tertulis diacu didalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, Juni 2023



Brigitha Pricilia Rahmi
NIM.1611016220005

ABSTRAK

PERBANDINGAN *SUPPORT VECTOR MACHINE*, *RANDOM FOREST* DAN *NAÏVE BAYES* PADA KLASIFIKASI SUARA BATUK COVID-19 (Oleh : Brigitha Pricilia Rahmi; Pembimbing : Mohammad Reza Faisal, S.T., M.T., Ph.D. dan Dwi Kartini, S.Kom., M.Kom.; 2023; 66 halaman)

COVID-19 (*Corona Virus Disease 2019*) merupakan penyakit menular yang menyerang saluran pernafasan atas yang disebabkan oleh virus corona dan ditetapkan sebagai pandemi global oleh World Health Organisation (WHO) pada 11 Maret 2020. Salah satu gejala yang ditimbulkan dari penyakit COVID-19 adalah batuk dan sesak nafas. Suara batuk yang dikeluarkan manusia beragam tergantung dari penyakit yang dideritanya, hal ini dikarenakan efek yang ditimbulkan penyakit terhadap saluran pernapasan manusia juga berbagai macam. Data masalah pernapasan seperti data suara batuk inilah yang kemudian dapat diproses dengan algoritma *machine learning* untuk mendiagnosa penyakit saluran pernafasan, termasuk penyakit COVID-19. Oleh karenanya peneliti mengusulkan penelitian untuk membandingkan metode klasifikasi *Support Vector Machine*, *Random Forest* dan *Naïve Bayes* dengan menggunakan dataset spektogram audio batuk COVID-19. Hasil evaluasi kinerja AUC dari masing-masing klasifikasi meliputi: *Support Vector Machine* dengan kernel *radial basis function* menghasilkan 0,5413 AUC, *Support Vector Machine* dengan kernel *polynomial* menghasilkan 0,5269 AUC, *Support Vector Machine* dengan kernel *linear* menghasilkan 0,5086 AUC, *Random Forest* menghasilkan 0,5545, serta *Naïve Bayes* menghasilkan 0,5807 AUC.

Kata Kunci : COVID-19, *Naïve Bayes*, *Random Forest*, *Support Vector Machine*, Suara Batuk

ABSTRACT

COMPARISON OF SUPPORT VECTOR MACHINE, RANDOM FOREST AND NAÏVE BAYES ON CLASSIFICATION OF COVID-19 COUGH SOUNDS

(By : Brigitha Pricilia Rahmi; Supervisors : Mohammad Reza Faisal, S.T., M.T., Ph.D. and Dwi Kartini, S.Kom., M.Kom.; 2023; 66 pages)

COVID-19 (Corona Virus Disease 2019) is an infectious disease that attacks the upper respiratory tract caused by a corona virus and was declared a global pandemic by the World Health Organization (WHO) on March 11, 2020. One of the symptoms caused by COVID-19 is cough and shortness of breath. The sound of coughing that humans make varies depending on the disease they are suffering from, this is because the effects that disease causes on the human respiratory tract also vary. Respiratory problem data such as cough sound data can then be processed with machine learning algorithms to diagnose respiratory tract diseases, including COVID-19 disease. Therefore this research proposed a study to compare the *Support Vector Machine*, *Random Forest* and *Naïve Bayes* classification methods using the COVID-19 cough audio spectrogram dataset. The results of the AUC performance evaluation of each classification include: *Support Vector Machine* with a *radial basis function* kernel resulted in 0.5413 AUC, *Support Vector Machine* with *polynomial* kernel resulted in 0.5269 AUC, *Support Vector Machine* with *linear* kernel produces 0.5086 AUC, *Random Forest* produces 0.5545, and *Naïve Bayes* resulted in 0.5807 AUC.

Keywords : Cough Sounds, COVID-19, *Naïve Bayes*, *Random Forest*, *Support Vector Machine*

PRAKATA

Assalamualaikum Warahmatullahi Wabarakatuh.

Puji syukur penulis panjatkan kepada Allah SWT karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul **Perbandingan Support Vector Machine, Random Forest dan Naïve Bayes pada Klasifikasi Suara Batuk COVID-19** untuk memenuhi syarat dalam menyelesaikan pendidikan program S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat. Tak lupa pula penulis panjatkan sholawat dan salam ke hadirat Rasulullah Muhammad SAW beserta para sahabat, keluarga, dan pengikut beliau hingga yaumul qiama.

Pada lembar ini penulis ingin menyampaikan ucapan terimakasih kepada pihak-pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan skripsi ini, diantaranya:

1. Keluarga terutama kedua orang tua saya bapak Bachruddin dan Ibu Misyati yang selalu memberikan dukungan, doa, dan bantuan dalam proses penyelesaian skripsi ini.
2. Bapak Mohammad Reza Faisal, S.T., M.T., Ph.D. dan Ibu Dwi Kartini, S.Kom, M.Kom. selaku dosen pembimbing yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
3. Bapak Dodon Turianto Nugrahadi, S.Kom., M.Eng. selaku ketua penguji serta bapak Muhammad Itqan Mazdadi, S.Kom., M.Kom. selaku anggota penguji yang banyak memberi saran dan meluangkan waktu untuk perbaikan skripsi ini.
4. Seluruh dosen dan staf Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
5. Teman-teman terdekat saya dan Team SVT yang telah bersedia membantu memecahkan kebingungan-kebingungan saya dan memberikan saran, dukungan, semangat, serta motivasi, selama proses penyelesaian skripsi.
6. Teman-teman keluarga Ilmu Komputer angkatan 2016 yang memberikan dukungan dan bantuan semasa kuliah dan selama proses penyelesaian skripsi.
7. Semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam penyelesaian skripsi ini.

Akhir kata penulis menyadari sepenuhnya bahwa penulisan ini jauh dari sempurna, namun penulis mengharapkan bantuan berupa saran dan kritik yang membangun dari semua pihak demi kesempurnaan dan mutu penulisan skripsi ini. Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, Juni 2023

A handwritten signature in black ink, appearing to read 'Brigitha Pricilia Rahmi'.

Brigitha Pricilia Rahmi

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	ii
PERNYATAAN.....	iii
ABSTRAK	iv
ABSTRACT	v
PRAKATA.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan.....	2
1.4 Manfaat Penelitian.....	3
1.5 Batasan Masalah.....	3
BAB II TINJAUAN PUSTAKA.....	4
2.1 Kajian Terdahulu	4
2.2 Keaslian Penelitan	7
2.3 COVID-19.....	12
2.4 Data Mining.....	12
2.5 Klasifikasi.....	13
2.6 Normalisasi Min-Max	14
2.7 <i>Cross Validation</i>	15
2.8 <i>Support Vector Machine</i>	16
2.9 <i>Random Forest</i>	17
2.10 <i>Naïve Bayes</i>	21
2.11 <i>Confusion Matrix</i>	22
2.13 <i>Area Under the ROC (Receiver Operating Characteristic) Curve</i>	23
BAB III METODE PENELITIAN	25

3.1	Alat Penelitian	25
3.2	Bahan Penelitian.....	25
3.3	Prosedur Penelitian.....	25
BAB IV HASIL DAN PEMBAHASAN		30
4.1	Hasil.....	30
4.1.1	Pengumpulan Data	30
4.1.2	Normalisasi	32
4.1.3	<i>Cross Validation</i>	34
4.1.4	Klasifikasi	35
4.2	Pembahasan	44
BAB V PENUTUP.....		48
5.1	Kesimpulan.....	48
5.2	Saran.....	48
DAFTAR PUSTAKA		49
LAMPIRAN.....		53
RIWAYAT HIDUP		67

DAFTAR TABEL

Tabel	Halaman
Tabel 1. Keaslian Penelitian.....	8
Tabel 2. Perancangan Penelitian	11
Tabel 3. Kernel pada SVM.....	17
Tabel 4. <i>Confusion Matrix</i>	22
Tabel 5. Keterangan rumus <i>Confusion matrix</i>	23
Tabel 6. Data Spektogram Audio Batuk COVID-19	26
Tabel 7. Data untuk fitur chroma_stft kelas COVID-19.....	27
Tabel 8. Data asli.....	30
Tabel 9. Keterangan fitur dataset	32
Tabel 10. Normalisasi data.....	32
Tabel 11. 10 K-fold cross validation.....	35
Tabel 12. <i>Confusion matrix</i> untuk klasifikasi SVM dengan kernel RBF	35
Tabel 13. Perbandingan <i>hyperparameter</i> c dan gamma kernel RBF	36
Tabel 14. Evaluasi per <i>fold</i> untuk klasifikasi SVM kernel RBF.....	36
Tabel 15. Hasil kinerja AUC klasifikasi SVM dengan kernel RBF	37
Tabel 16. <i>Confusion matrix</i> untuk klasifikasi SVM dengan kernel <i>polynomial</i>	37
Tabel 17. Perbandingan <i>hyperparameter</i> c dan gamma kernel <i>polynomial</i>	38
Tabel 18. Evaluasi per <i>fold</i> untuk klasifikasi SVM kernel <i>Polynomial</i>	38
Tabel 19. Hasil kinerja AUC klasifikasi SVM dengan kernel <i>polynomial</i>	39
Tabel 20. <i>Confusion matrix</i> untuk klasifikasi SVM dengan kernel <i>linear</i>	39
Tabel 21. Perbandingan <i>hyperparameter</i> c dan gamma kernel <i>linear</i>	40
Tabel 22. Evaluasi per <i>fold</i> untuk klasifikasi SVM kernel <i>linear</i>	40
Tabel 23. Hasil kinerja AUC klasifikasi SVM dengan kernel <i>linear</i>	41
Tabel 24. <i>Confusion matrix</i> untuk klasifikasi <i>Random Forest</i>	41
Tabel 25. Perbandingan <i>hyperparameter</i> pada <i>Random Forest</i>	42
Tabel 26. Evaluasi per <i>fold</i> untuk klasifikasi <i>Random Forest</i>	42
Tabel 27. Hasil kinerja AUC klasifikasi <i>Random Forest</i>	43
Tabel 28. <i>Confusion matrix</i> untuk klasifikasi <i>Naïve Bayes</i>	43
Tabel 29. Hasil kinerja AUC klasifikasi <i>Naïve Bayes</i>	43

DAFTAR GAMBAR

Gambar	Halaman
Gambar 1. Penjelasan 10 <i>Folds Cross Validation</i>	16
Gambar 2. Data non-linear	17
Gambar 3. Contoh <i>Random Forest</i>	18
Gambar 4. Contoh <i>Random Forest</i>	19
Gambar 5. Nilai AUC	24
Gambar 6. Alur penelitian.....	26
Gambar 7. Data untuk kelas COVID-19 pada fitur <i>chroma_stft</i>	27
Gambar 8. Visualisasi kurva untuk nilai AUC klasifikasi SVM kernel RBF	44
Gambar 9. Visualisasi kurva untuk nilai AUC klasifikasi SVM kernel <i>polynomial</i> .	45
Gambar 10. Visualisasi kurva untuk nilai AUC klasifikasi SVM kernel <i>linear</i>	45
Gambar 11. Visualisasi kurva untuk nilai AUC klasifikasi <i>Random Forest</i>	46
Gambar 12. Visualisasi kurva untuk nilai AUC klasifikasi <i>Naïve Bayes</i>	46
Gambar 13. Grafik perbandingan rata-rata AUC.....	47

DAFTAR LAMPIRAN

Lampiran	Halaman
Lampiran 1. <i>Source code</i> normalisasi data	54
Lampiran 2. <i>Source code</i> klasifikasi <i>Support Vector Machine</i> kernel RBF	54
Lampiran 3. <i>Source code</i> klasifikasi <i>Support Vector Machine</i> kernel <i>polynomial</i>	55
Lampiran 4. <i>Source code</i> klasifikasi <i>Support Vector Machine</i> kernel <i>linear</i>	56
Lampiran 5. <i>Source code</i> klasifikasi <i>Random Forest</i>	56
Lampiran 6. <i>Source code</i> klasifikasi <i>Naïve Bayes</i>	57
Lampiran 7. Hasil perbandingan <i>hyperparameter</i> untuk klasifikasi <i>Support Vector Machine</i> dengan kernel <i>radial basis function</i> (RBF).....	57
Lampiran 8. Hasil perbandingan <i>hyperparameter</i> untuk klasifikasi <i>Support Vector Machine</i> dengan kernel <i>polynomial</i>	59
Lampiran 9. Hasil perbandingan <i>hyperparameter</i> untuk klasifikasi <i>Support Vector Machine</i> dengan kernel <i>linear</i>	60
Lampiran 10. Hasil perbandingan <i>hyperparameter</i> untuk klasifikasi <i>Random Forest</i>	61