

## Ensemble Learning dengan Soft Voting Classifier untuk Klasifikasi Pasien Tifus di Puskesmas Balung

Lutfi Ali Muharom<sup>1</sup>, Dudi Irawan<sup>2</sup>, Taufiq Timur Warisaji<sup>3</sup>

<sup>1</sup>Teknik Informatika, Universitas Muhammadiyah Jember, [lutfi.muharom@unmuhjember.ac.id](mailto:lutfi.muharom@unmuhjember.ac.id)

<sup>2</sup>Teknik Informatika, Universitas Muhammadiyah Jember, [dudi.irawan@unmuhjember.ac.id](mailto:dudi.irawan@unmuhjember.ac.id)

<sup>3</sup>Teknik Informatika, Universitas Muhammadiyah Jember, [taufiqtimur@unmuhjember.ac.id](mailto:taufiqtimur@unmuhjember.ac.id)

### Keywords

Ensemble Learning  
Gradient Boosting  
Logistic Regression  
Random Forest  
Typhoid Fever

### ABSTRACT

Typhoid fever remains a significant public health problem in Indonesia, particularly in areas with limited medical facilities. This study aims to develop an automatic classification model for typhoid diagnosis using an ensemble learning approach based on the Soft Voting Classifier. The model combines three base algorithms, Logistic Regression, Random Forest, and Gradient Boosting, to enhance predictive accuracy. The dataset was obtained from Balung Primary Health Center, Jember Regency, consisting of 510 patient records with typhoid symptoms. Experimental results show that the ensemble model achieved an accuracy of over 92%, outperforming individual models. Furthermore, adequate precision and recall indicate the model's potential to support rapid and accurate medical diagnosis. These findings demonstrate that the Soft Voting Classifier can serve as an effective tool to assist healthcare workers, especially in resource-limited settings, in improving the quality of typhoid fever diagnosis.

### Kata Kunci

Demam Tifoid  
Ensemble Learning  
Gradient Boosting  
Logistic Regression  
Random Forest

### ABSTRAK

Demam tifoid masih menjadi salah satu masalah kesehatan serius di Indonesia, terutama di wilayah dengan keterbatasan fasilitas medis. Penelitian ini bertujuan mengembangkan model klasifikasi otomatis untuk diagnosis tifus menggunakan pendekatan ensemble learning berbasis Soft Voting Classifier. Model ini menggabungkan tiga algoritma dasar, yaitu Logistic Regression, Random Forest, dan Gradient Boosting, guna meningkatkan akurasi prediksi. Dataset yang digunakan berasal dari Puskesmas Balung Kabupaten Jember dengan total 510 data pasien yang menunjukkan gejala tifus. Hasil pengujian menunjukkan bahwa model ensemble mampu mencapai akurasi lebih dari 92%, lebih tinggi dibandingkan penggunaan model tunggal. Selain itu, precision dan recall yang memadai menunjukkan potensi model ini untuk mendukung diagnosis medis yang cepat dan akurat. Temuan ini memperlihatkan bahwa penggunaan Soft Voting Classifier dapat menjadi solusi efektif dalam membantu tenaga kesehatan, khususnya di daerah dengan keterbatasan sumber daya, untuk meningkatkan kualitas diagnosis demam tifoid.

### Korespondensi Penulis:

Lutfi Ali Muharom  
Universitas Muhammadiyah Jember, Jl. Karimata 49 Jember  
Telepon : +6281357144055  
Email: [lutfi.muharom@unmuhjember.ac.id](mailto:lutfi.muharom@unmuhjember.ac.id)

Submitted : 25-09-2025; Accepted : 30-09-2025;

Published : 30-09-2029

*Copyright (c) 2025 The Author (s) This article is distributed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0)*

## 1. PENDAHULUAN

Demam tifoid atau umum disebut penyakit tifus adalah demam yang berlangsung cukup lama dengan gejala demam tinggi, sembelit, sakit perut, pusing, gatal dan bintik kemerahan akibat bakteremia disertai peradangan yang merusak usus dan hati. Demam tifoid merupakan penyakit menular yang ada di seluruh dunia dan merupakan masalah kesehatan terbesar di negara-negara berkembang seperti Asia Tenggara, Afrika dan Amerika Latin. Insiden demam tifoid masih sangat tinggi, diperkirakan dari 21 juta kasus, ada 200.000 di antaranya meninggal setiap tahun. Kasus demam tifoid di Indonesia merupakan masalah kesehatan masyarakat endemis yang dari tinjauan kasus di rumah sakit besar Indonesia cenderung meningkat setiap tahun dengan rata-rata kasus 500 per 100.000 penduduk dengan tingkat kematian antara 0,6% sampai 5% [1]. Di Indonesia, tifus masih menjadi salah satu masalah kesehatan yang cukup serius, terutama di daerah pedesaan dan wilayah dengan akses sanitasi yang terbatas [2].

Penyakit tifus, yang disebabkan oleh *Salmonella enterica serovar Typhi*, umumnya memerlukan diagnosis yang tepat dan cepat untuk mengurangi tingkat komplikasi dan kematian [3]. Di Indonesia, di mana kasus demam tifoid masih cukup tinggi, diagnosis yang tepat seringkali terhambat oleh keterbatasan fasilitas kesehatan, terutama di daerah pedesaan [4]. Salah satu tantangan utama dalam penanganan penyakit ini adalah mengidentifikasi pasien yang terinfeksi berdasarkan gejala klinis yang mirip dengan banyak penyakit lainnya [5]. Oleh karena itu, pendekatan berbasis teknologi, khususnya dalam bidang kecerdasan buatan dan pembelajaran mesin, menawarkan peluang besar untuk mempercepat proses klasifikasi dan diagnosis penyakit tifus secara otomatis, menggunakan data medis yang ada [6], [7], [8], [9].

Berbagai metode pembelajaran mesin dan kecerdasan buatan telah banyak diaplikasikan dalam bidang medis untuk membantu diagnosis penyakit, mulai dari model sederhana seperti regresi logistik hingga model yang lebih kompleks seperti pohon keputusan, *support vector machine*, maupun *boosting* [10], [11], [12], [13]. Masing-masing metode tersebut memiliki performa yang bervariasi tergantung pada kualitas data, jumlah fitur, serta karakteristik kasus medis yang dianalisis. Namun, penggunaan satu model tunggal sering kali belum cukup konsisten dalam memberikan hasil yang optimal. Oleh karena itu, pendekatan *ensemble* yang menggabungkan beberapa model sekaligus menjadi solusi yang lebih menjanjikan, karena mampu memanfaatkan kelebihan dari setiap algoritma sekaligus meminimalkan kelemahan yang dimiliki masing-masing model [14].

Penerapan Model Ensemble *Soft Voting Classifier* yang menggabungkan *Logistic Regression*, *Random Forest*, dan *Gradient Boosting* dapat meningkatkan akurasi dalam klasifikasi bidang medis. *Logistic Regression* efektif untuk menangani hubungan linier antar fitur, *Random Forest* mampu memproses data kompleks melalui ensemble pohon keputusan, sementara *Gradient Boosting* unggul dalam menghadapi *outlier* dan data tidak seimbang dengan prediksi iteratif yang presisi [15], [16], [17], [18]. Dengan menggabungkan kekuatan ketiga algoritma tersebut melalui *soft voting*, sistem dapat menghasilkan keputusan klasifikasi yang lebih akurat, stabil, dan andal [15], [17], [18]. Pendekatan ini berpotensi mendukung diagnosis tifus secara otomatis melalui analisis data medis seperti gejala, hasil laboratorium, dan riwayat pasien, sehingga membantu tenaga medis membuat keputusan cepat dan tepat, khususnya di wilayah dengan keterbatasan sumber daya.

Dalam penelitian ini, akan digunakan model *Soft Voting Classifier* yang menggabungkan prediksi dari tiga algoritma dasar, yaitu *Logistic Regression*, *Random Forest*, dan *Gradient Boosting*. Model ini diharapkan dapat menghasilkan prediksi yang lebih akurat mengenai apakah seorang pasien positif atau negatif tifus berdasarkan atribut-atribut klinis yang ada. Data yang tersedia dari Puskesmas Kecamatan Balung Kabupaten Jember mencatat adanya 510 kasus pasien yang dirawat dengan gejala tifus, di mana 421 pasien terdiagnosis negatif tifus dan 89 pasien terdiagnosis positif tifus. Mendeteksi tifus secara dini melalui gejala klinis dan pemeriksaan laboratorium sangat penting untuk penanganan yang tepat.

## 2. METODE PENELITIAN

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan *machine learning* untuk menganalisis dataset medis yang berisi informasi tentang pasien yang terdiagnosis dengan gejala demam tifoid. Penelitian ini bertujuan untuk mengembangkan model klasifikasi berbasis ensemble learning untuk membantu diagnosis tifus secara otomatis berdasarkan atribut klinis yang tersedia. Model yang akan digunakan adalah *Soft Voting Classifier*, yang menggabungkan tiga algoritma dasar, yaitu *Logistic Regression*, *Random Forest*, dan *Gradient Boosting*.

Tahapan penelitian meliputi Pengumpulan Data, *Preprocessing Data*, Klasifikasi, dan Evaluasi.



Gambar 1. Metodologi Penelitian

Data yang digunakan dalam penelitian ini diperoleh dari Puskesmas Kecamatan Balung Kabupaten Jember, yang mencatat 510 kasus pasien yang dirawat dengan gejala tifus. Atribut yang dicatat mengambil acuan dari penelitian Devita dan kawan – kawan [19] Dataset ini terdiri dari 6 atribut yang meliputi ‘Demam tinggi’ merupakan indikator demam pada pasien, ‘Sakit kepala’ menjelaskan apakah pasien mengalami sakit kepala, Badan lemas menjelaskan apakah pasien merasa lemas atau tidak, ‘Sakit perut’ adalah kondisi apakah ada nyeri pada perut, ‘Penurunan berat badan’ apakah pasien mengalami penurunan berat badan yang signifikan, ‘Hasil Tifus’ adalah label yang menunjukkan apakah pasien terdiagnosis positif atau negatif tifus, dengan dua label: Negatif dan Positif.

*Preprocessing data* dilakukan sebelum model dibangun menggunakan model pembelajaran mesin, *preprocessing data* merupakan proses persiapan data, Pembersihan Data dan Perubahan Data. Pembersihan data menghapus data yang hilang atau tidak lengkap. Perubahan Data adalah merubah nilai pada atribut agar siap untuk proses selanjutnya.

*Ensemble classification* dalam penelitian ini akan menggunakan model *Soft Voting Classifier*, yang merupakan teknik ensemble yang menggabungkan hasil prediksi dari tiga algoritma dasar, yaitu:

1. Logistic Regression: Digunakan untuk menangani hubungan linier antar fitur dan memberikan dasar pemahaman terhadap hubungan antara variabel-variabel yang ada. Secara umum rumus dari Logistic Regression adalah :

$$\frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i)}} \quad (1)$$

dimana  $\beta_0, \beta_1, \beta_i$  adalah koefisien dari persamaan regresi dan  $x_1, x_2, x_i$  adalah variabel independen.

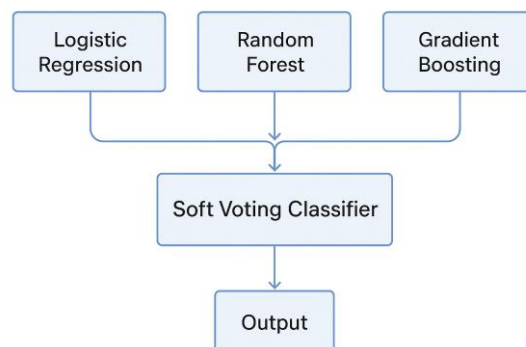
2. *Random Forest*: Digunakan untuk menangani data yang lebih kompleks dengan menggunakan banyak pohon keputusan yang mengurangi risiko *overfitting*.

Cara Kerja Algoritma *Random Forest*:

- Pertama algoritma ini membentuk decision tree atau pohon keputusan, decision tree yang dibentuk lebih dari 1 dan masing – masing tree atau pohon memilih fitur pada dataset secara acak, sehingga setiap pohon berbeda satu sama lain.
- Kemudian setiap pohon memberikan jawaban atau prediksinya sendiri berdasarkan apa yang dipelajarinya dari bagian datanya.
- Tahap terakhir adalah menggabungkan hasil prediksi, sehingga proses klasifikasi dipilih berdasarkan mayoritas dari hasil setiap pohon.

3. Gradient Boosting: Dikenal dengan kemampuannya untuk mengatasi masalah data tidak seimbang dan mampu memberikan prediksi yang presisi dengan meningkatkan model secara bertahap melalui teknik boosting.

Proses *ensemble* dilakukan dengan cara *soft voting*, di mana setiap model mengeluarkan prediksi probabilitas dan hasilnya dihitung berdasarkan rata-rata probabilitas dari setiap algoritma untuk menentukan prediksi akhir. Pendekatan ini diharapkan dapat meningkatkan akurasi, stabilitas, dan reliabilitas dalam proses diagnosis. Proses *ensemble* ditunjukkan pada gambar 2.



Gambar 2. Proses *Ensemble*

Untuk mengevaluasi performa model yang dikembangkan, beberapa metrik evaluasi yang akan digunakan antara lain:

- Akurasi: Untuk mengukur seberapa banyak prediksi yang benar dibandingkan dengan total jumlah data.
- *Precision, Recall, dan F1-Score*: Digunakan untuk menilai kemampuan model dalam mengidentifikasi pasien yang terdiagnosis positif tifus.

Confusion Matrix: Untuk memberikan gambaran lebih rinci tentang distribusi prediksi benar dan salah antara kelas Positif dan Negatif. Dalam Confusion Matrix terdapat 4 istilah representasi klasifikasi yaitu *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)*, *False Negatif (FN)*. Berikut adalah tabel representasi hasil klasifikasi sebagai berikut :

Tabel 1. *Confussion Matrix*

Correct classification	Classified as	
	Predicted “positive”	Predicted “negative”
Actual “positive”	True positives	False negatives
Actual “negative”	False positives	True negatives

Berdasarkan pada tabel 1 dapat diartikan sebagai berikut :

- a) *True Positive (TP)* adalah data positif terjangkit tifus baik pada prediksi maupun aktual

- b) *True Negative* (TN) adalah data negatif terjangkau tifus baik pada prediksi maupun aktual
  - c) *False Positive* (FP) adalah ketika data secara aktual negatif tifus tetapi hasil prediksi dinyatakan positif
  - d) *False Negatif* (FN) adalah ketika data secara aktual positif tifus tetapi hasil prediksi dinyatakan negatif
- Rumus dari tiap peformasi adalah :

$$\text{Accuracy (Akurasi)} = \frac{TP+TN}{TP+TN+FN+FP} \quad (1)$$

$$\text{Precision (Presisi)} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall (Sensitivitas)} = \frac{TP}{TP+FN} \quad (3)$$

### 3. HASIL DAN ANALISIS

Setelah melakukan pelatihan pada dataset yang telah dibersihkan, model *Soft Voting Classifier* diuji untuk melihat kinerjanya dalam klasifikasi pasien tifus. Evaluasi model dilakukan dengan menggunakan metrik-metrik evaluasi seperti akurasi, *precision*, *recall*, dan *F1-score* untuk menilai kinerja dalam mengklasifikasikan pasien sebagai positif atau negatif tifus. Berdasarkan hasil percobaan, model *Soft Voting Classifier* menunjukkan performa yang lebih baik dibandingkan dengan masing-masing model dasar secara individual, baik dalam hal akurasi maupun metrik evaluasi lainnya.

#### 3.1 Pengumpulan Data

Dalam penelitian ini, dataset yang digunakan berasal dari Puskesmas Balung, data anamnesis pasien yang berkunjung ke Puskesmas Balung, terdiri dari data pengukuran suhu badan, data kondisi sakit kepala, data kondisi badan, keluhan sakit perut, kondisi penurunan berat badan, dan diagnosis hasil tifus (positif atau negatif).

Tabel 2. Data Pasien

Suhu	Sakit kepala	Badan lemas	Sakit perut	Penurunan berat badan	Hasil Tifus
36	Ya	Tidak	Tidak	Tidak	Negatif
36	Ya	Ya	Tidak	Tidak	Negatif
37	Ya	Ya	Tidak	Tidak	Negatif
35	Ya	Ya	Tidak	Tidak	Negatif
37	Ya	Ya	Tidak	Turun	Positif
35	Tidak	Tidak	Tidak	Tidak	Negatif
34	Tidak	Tidak	Tidak	Turun	Negatif
34	Tidak	Ya	Tidak	Turun	Positif
34	Ya	Tidak	Tidak	Tidak	Negatif
34	Tidak	Ya	Ya	Tidak	Positif
33	Tidak	Ya	Ya	Tidak	Positif
33	Tidak	Ya	Ya	Tidak	Negatif

Data pengukuran suhu tubuh menggunakan satuan derajat celsius. Data tersebut adalah para pasien yang berkunjung ke Puskesmas Balung selama tahun 2019.

#### 3.2 Preprocessing Data

Pada tahap preprocessing data, yang dilakukan adalah pembersihan data, konversi atau perubahan data, dan pelabelan. Proses pembersihan data dilakukan pada dataset untuk menghapus baris data yang tidak lengkap, kriteria baris data yang dihapus adalah jika pada suatu baris data ada yang tidak memiliki nilai pada salah satu kolom atribut, maka baris data tersebut dihapus agar tidak menyebabkan salah perhitungan. Setelah dilakukan proses pembersihan data, dari total 621 data maka jumlah berkurang menjadi 510 data.

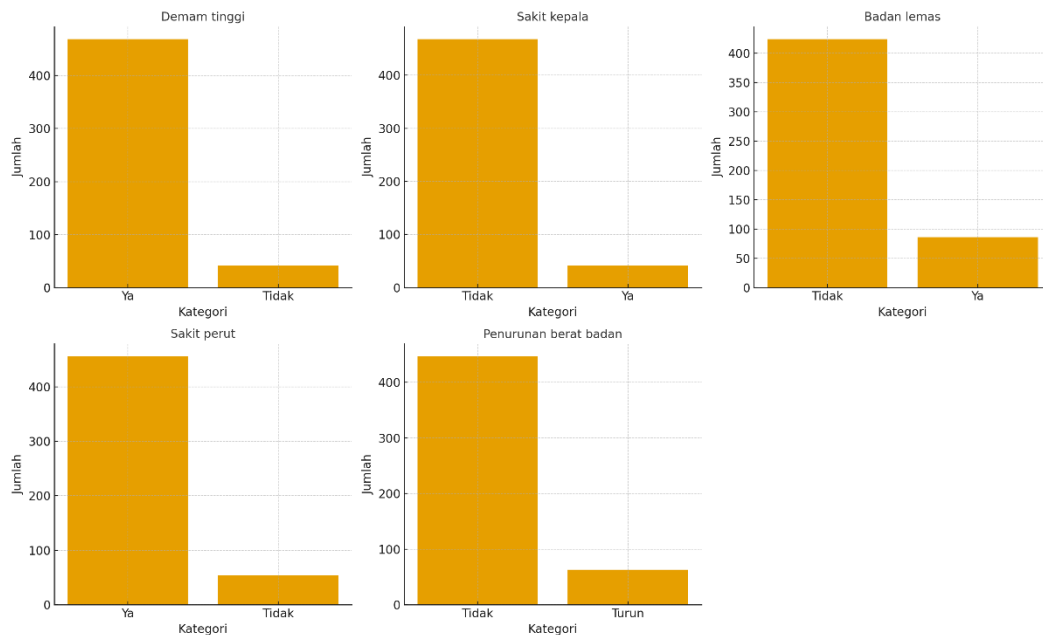
Pada tahap perubahan data, perubahan nama atribut ‘Suhu’ menjadi ‘Demam tinggi’ data dan perubahan angka suhu tubuh kedalam jenis data nominal. sesuai penelitian medis oleh Breda [20], jika suhu tubuh diatas 30 derajat celsius maka diganti dengan ‘Ya’ dan apabila suhu tubuh 30 derajat celsius kebawah maka diganti dengan ‘Tidak’, sesuai penelitian medis oleh Breda dan kawan – kawan. Terakhir adalah menambahkan sebuah atribut atau kolom ‘Hasil Tifus’ yang berisi tentang label kelas dari baris data atau pasien yang bersangkutan, apakah terjangkau tifus atau tidak, jika terjangkau maka diisi dengan Positif, dan jika tidak terjangkau maka diisi dengan ‘Negatif’. Label data ini dilakukan dan divalidasi oleh dokter puskesmas.

Tabel 3 menunjukkan dataset setelah dilakukan preprocessing, dan gambar 2 menunjukkan distribusi kategori dari setiap atribut dataset.

Tabel 3. Dataset Setelah Preprocessing

Demam tinggi	Sakit kepala	Badan lemas	Sakit perut	Penurunan berat badan	Hasil Tifus
Ya	Ya	Tidak	Tidak	Tidak	Negatif

Ya	Ya	Ya	Tidak	Tidak	Negatif
Ya	Ya	Ya	Tidak	Tidak	Negatif
Ya	Ya	Ya	Tidak	Tidak	Negatif
Ya	Ya	Ya	Tidak	turun	Positif
Ya	Tidak	Tidak	Tidak	Tidak	Negatif
Ya	Tidak	Tidak	Tidak	Turun	Negatif
Ya	Tidak	Ya	Tidak	Turun	Positif
Ya	Ya	Tidak	Tidak	Tidak	Negatif
Ya	Tidak	Ya	Ya	Tidak	Positif
Ya	Tidak	Ya	Ya	Tidak	Positif



Gambar 3. Distribusi Nilai Atribut

### 3.3 Klasifikasi

Setelah melakukan pelatihan pada dataset yang telah dibersihkan, model *Soft Voting Classifier* diuji untuk melihat kinerjanya dalam klasifikasi pasien tifus. Evaluasi model ensemble ini mempunyai nilai akurasi mencapai 97% ketika memakai pembagian data 80% data latih dan 20% data uji, yang menunjukkan bahwa model ini dapat melakukan klasifikasi dengan baik.

Pada tahap klasifikasi, *Logistik Regression* menggunakan parameter iterasi 1000 dan random state 42, untuk Random Forest menggunakan 100 *estimator* dan *random state* 42, dan pada *Gradient Boosting* menggunakan estimator 100 dan random state 42. Sedangkan untuk *cross validation*, pada penelitian ini menggunakan 2 sampai 10 *fold cross validation*.

Tabel 4. Ringkasan Rata-rata Akurasi

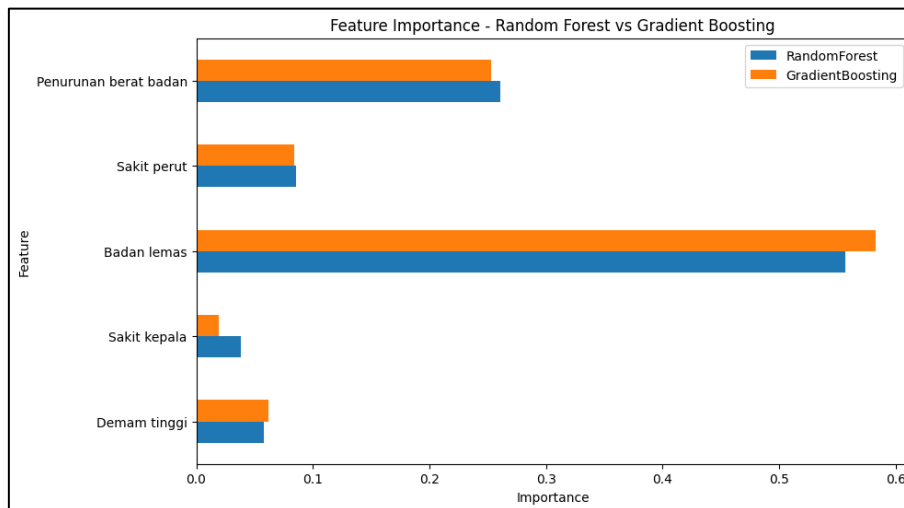
Fold	Rata-rata Akurasi
2	88,0392%
3	87,8431%
4	87,6538%
5	87,4510%
6	92,5490%
7	92,3516%
8	92,1441%
9	92,1436%
10	86,4706%

Akurasi model *Soft Voting* mencapai sekitar 92%, yang menunjukkan bahwa model ini berhasil mengklasifikasikan lebih dari 92% pasien dengan benar. Precision dan recall masing-masing menunjukkan angka yang cukup memadai, mengindikasikan bahwa model ini mampu mengenali dengan baik pasien yang benar-benar

positif tifus tanpa banyak menghasilkan prediksi positif palsu. Dengan demikian, model ini memberikan potensi yang baik dalam mendukung keputusan medis di Puskesmas Balung dalam mengidentifikasi pasien tifus secara lebih akurat.

### 3.4 Evaluasi Fitur

Evaluasi fitur dilakukan untuk mengetahui fitur atau faktor diagnosis manakah yang berpengaruh terhadap penyakit tifus.



Gambar 4. Diagram Feature Importance

Batang berwarna biru menunjukkan tingkat kepentingan menurut model *Random Forest*, sedangkan batang berwarna oranye menunjukkan tingkat kepentingan menurut model *Gradient Boosting*. Panjang setiap batang menunjukkan seberapa besar kontribusi fitur spesifik tersebut terhadap prediksi model. Batang yang lebih panjang berarti fitur tersebut lebih penting. Dengan membandingkan panjang batang untuk setiap fitur, dapat diketahui gejala mana yang dianggap paling berpengaruh dalam memprediksi "Hasil Tifus" oleh masing – masing model. Secara grafik baik *Gradient Boosting* dan *Random Forest* sama – sama menilai bahwa atribut ‘Badan Lemas’ merupakan fitur paling berpengaruh dalam penyakit tifus, dan fitur ‘Sakit Kepala’ merupakan yang paling lemah mempengaruhi. Urutan pengaruh masing – masing fitur tidak berbeda secara urutan, hanya nilai pengaruh yang sedikit berbeda.

## 4. KESIMPULAN

Penelitian ini telah menunjukkan bahwa penggunaan model *Soft Voting Classifier* yang menggabungkan tiga algoritma dasar *Logistic Regression*, *Random Forest*, dan *Gradient Boosting* dapat meningkatkan akurasi dalam klasifikasi pasien tifus berdasarkan data atribut klinis. Model ini memiliki potensi untuk menjadi alat yang berguna dalam mendukung diagnosis tifus di tingkat Puskesmas, di mana akurasi diagnosis dapat meningkatkan kualitas perawatan pasien dan mengurangi kesalahan diagnosa.

Ke depan, penelitian lebih lanjut dapat dilakukan dengan menguji model ini pada dataset yang lebih besar atau mempertimbangkan variabel tambahan yang dapat mempengaruhi diagnosis tifus, seperti data laboratorium atau faktor epidemiologis lainnya. Dengan pengembangan lebih lanjut, model ini dapat diintegrasikan ke dalam sistem kesehatan untuk membantu petugas medis dalam pengambilan keputusan klinis yang lebih cepat dan akurat.

## REFERENSI

- [1] Budayatin, J. Waluyo, and D. Wahyuni, "The identification of Salmonella sp. in 'cilok' road food in campus area of Jember University," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 747, no. 1, p. 12106, 2021, doi: 10.1088/1755-1315/747/1/012106.
- [2] K. Saraswati, J. K. Baird, S. D. Blacksell, M. L. Grijnsen, and N. P. J. Day, "History of scrub typhus in Indonesia," *Trans. R. Soc. Trop. Med. Hyg.*, vol. 119, no. 4, pp. 338–345, Apr. 2025, doi: 10.1093/trstmh/traf017.
- [3] D. P. Neupane, H. P. Dulal, and J. Song, "Enteric Fever Diagnosis: Current Challenges and Future Directions," *Pathogens*, vol. 10, no. 4. 2021. doi: 10.3390/pathogens10040410.
- [4] J. Sapkota, T. Roberts, B. Basnyat, S. Baker, L. M. Hampton, and S. Dittrich, "Diagnostics for Typhoid Fever: Current Perspectives and Future Outlooks for Product Development and Access," *Open Forum Infect. Dis.*, vol. 10, no. Supplement\_1, pp. S17–S20, May 2023, doi: 10.1093/ofid/ofad120.
- [5] K. Kannan *et al.*, "Performance of molecular and serologic tests for the diagnosis of scrub typhus," *PLoS Negl. Trop. Dis.*, vol. 14, no. 11, p. e0008747, Nov. 2020, [Online]. Available:

- <https://doi.org/10.1371/journal.pntd.0008747>
- [6] F. E. Ayo, J. B. Awotunde, R. O. Ogundokun, S. O. Folorunso, and A. O. Adekunle, "A decision support system for multi-target disease diagnosis: A bioinformatics approach," *Heliyon*, vol. 6, no. 3, Mar. 2020, doi: 10.1016/j.heliyon.2020.e03657.
- [7] S. Mishra, R. Kumar, S. K. Tiwari, and P. Ranjan, "Machine learning approaches in the diagnosis of infectious diseases: a review," *Bull. Electr. Eng. Informatics; Vol 11, No 6 December 2022DO - 10.11591/eei.v11i6.4225*, Dec. 2022, [Online]. Available: <https://beei.org/index.php/EEI/article/view/4225>
- [8] S. Roy, T. Meena, and S.-J. Lim, "Demystifying Supervised Learning in Healthcare 4.0: A New Reality of Transforming Diagnostic Medicine," *Diagnostics*, vol. 12, no. 10, 2022. doi: 10.3390/diagnostics12102549.
- [9] F. A. Nishat, M. F. Mridha, I. Mahmud, M. Alfarhood, M. Safran, and D. Che, "Enhancing Typhoid Fever Diagnosis Based on Clinical Data Using a Lightweight Machine Learning Metamodel," *Diagnostics*, vol. 15, no. 5, 2025. doi: 10.3390/diagnostics15050562.
- [10] J. Qin, L. Chen, Y. Liu, C. Liu, C. Feng, and B. Chen, "A Machine Learning Methodology for Diagnosing Chronic Kidney Disease," *IEEE Access*, vol. 8, pp. 20991–21002, 2020, doi: 10.1109/ACCESS.2019.2963053.
- [11] A. M. Rahmani *et al.*, "Machine Learning (ML) in Medicine: Review, Applications, and Challenges," *Mathematics*, vol. 9, no. 22, 2021. doi: 10.3390/math9222970.
- [12] N. Caballé-Cervigón, J. L. Castillo-Sequera, J. A. Gómez-Pulido, J. M. Gómez-Pulido, and M. L. Polo-Luque, "Machine Learning Applied to Diagnosis of Human Diseases: A Systematic Review," *Applied Sciences*, vol. 10, no. 15, 2020. doi: 10.3390/app10155135.
- [13] G. Battineni, G. G. Sagaro, N. Chinatalapudi, and F. Amenta, "Applications of Machine Learning Predictive Models in the Chronic Disease Diagnosis," *Journal of Personalized Medicine*, vol. 10, no. 2, 2020. doi: 10.3390/jpm10020021.
- [14] P. Pintelas and I. E. Livieris, "Special Issue on Ensemble Learning and Applications," *Algorithms*, vol. 13, no. 6, 2020. doi: 10.3390/a13060140.
- [15] K. Cao-Van, T. C. Minh, L. G. Minh, T. T. B. Quyen, and H. M. Tan, "Soft-Voting Ensemble Model: An Efficient Learning Approach for Predictive Prostate Cancer Risk," *Vietnam J. Comput. Sci.*, vol. 11, no. 04, pp. 531–552, Jul. 2024, doi: 10.1142/S2196888824500155.
- [16] N. Chandrasekhar and S. Peddakrishna, "Enhancing Heart Disease Prediction Accuracy through Machine Learning Techniques and Optimization," *Processes*, vol. 11, no. 4, 2023. doi: 10.3390/pr11041210.
- [17] A. Srinivas and J. P. Mosiganti, "A brain stroke detection model using soft voting based ensemble machine learning classifier," *Meas. Sensors*, vol. 29, p. 100871, 2023, doi: <https://doi.org/10.1016/j.measen.2023.100871>.
- [18] S. W. A. Sherazi, J.-W. Bae, and J. Y. Lee, "A soft voting ensemble classifier for early prediction and diagnosis of occurrences of major adverse cardiovascular events for STEMI and NSTEMI during 2-year follow-up in patients with acute coronary syndrome," *PLoS One*, vol. 16, no. 6, p. e0249338, Jun. 2021, [Online]. Available: <https://doi.org/10.1371/journal.pone.0249338>
- [19] N. Devita, M. S. Riski, R. Marufi, and U. A. Habibah, "Profil manifestasi klinis dan laboratorium pasien demam tifoid di Rumah Sakit PKU Bantul," *J. Kedokt. Univ. Palangka Raya*, vol. 11, no. 2 SE-Articles, Oct. 2023, doi: 10.37304/jkupr.v11i2.10753.
- [20] J. Breda, M. Springston, A. Mariakakis, and S. Patel, "FeverPhone: Accessible Core-Body Temperature Sensing for Fever Monitoring Using Commodity Smartphones," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 7, no. 1, Mar. 2023, doi: 10.1145/3580850.