Bulletin of Data Science
ISSN 2745-7966 (Media Online)
Vol 5, No 3, June 2026 | Page 170−178
DOI: 10.47065/bulletinds.v5i3.10142
https://ejurnal.seminar-id.com/index.php/bulletinds

Klasifikasi Multi-Label Hadis Terjemahan Shahih Bukhari Berdasarkan Kategori Anjuran, Larangan, dan Informasi Menggunakan Metode *Extreme Gradient Boosting*

Tasya Dwi Yanti, Fitra Kurnia*, Nazruddin Safaat Harahap, Iwan Iskandar, Reski Mai Candra

Fakultas Sains & Teknologi, Prodi Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia
Email: [1]12250123081@students.uin-suska.ac, [2,*]fitra.k@uin-suska.ac.id, [3]nazruddin.safaat@uin-suska.ac.id, [4]iwan.iskandar@uin-suska.ac.id, [5]reski.candra@uin-suska.ac.id
Email Penulis Korespondensi: fitra.k@uin-suska.ac.id

Abstrak−Hadis merupakan salah satu sumber ajaran Islam yang memuat anjuran, larangan, dan informasi sebagai pedoman bagi umat Islam. Banyaknya jumlah hadis menyebabkan proses klasifikasi secara manual menjadi kurang efisien, sehingga diperlukan metode klasifikasi otomatis yang mampu mengelompokkan hadis ke dalam lebih dari satu kategori. Penelitian ini bertujuan untuk melakukan klasifikasi Multi-label pada hadis terjemahan Shahih Bukhari menggunakan metode *Extreme Gradient Boosting* (XGBoost) dengan pendekatan *Binary Relevance*. Untuk meningkatkan performa model, dilakukan *hyperparameter tuning* menggunakan *GridSearchCV*. *Dataset* yang digunakan terdiri dari 7000 data hadis terjemahan Shahih Bukhari yang telah diberi label ke dalam kategori anjuran, larangan, dan informasi. Hasil penelitian menunjukkan bahwa model mampu mencapai nilai *accuracy* hingga 94,64%, *precision* hingga 95,17%, *recall* hingga 99,47%, dan F1-score hingga 97,24%. Penerapan *hyperparameter tuning* menggunakan *GridSearchCV* meningkatkan nilai *recall* dari 32,72% menjadi 35,66% serta F1-score dari 47,09% menjadi 50,00%, sekaligus menurunkan nilai *hamming loss* dari 9,19% menjadi 9,10%. Hasil tersebut menunjukkan bahwa kombinasi XGBoost, *Binary Relevance*, dan *GridSearchCV* efektif digunakan untuk klasifikasi Multi-label pada hadis terjemahan Shahih Bukhari.

Kata Kunci: Hadis, Klasifikasi Multi-Label; XGBoost; *Binary Relevance; GridSearchCV*

Abstract−Hadith is one of the primary sources of Islamic teachings, containing recommendations, prohibitions, and information that serve as guidance for Muslims. The large number of hadiths makes manual classification inefficient and time-consuming, creating the need for an automated classification method capable of assigning multiple categories to a single hadith. This study aims to perform multi-label classification on translated Sahih Bukhari hadiths using the Extreme Gradient Boosting (XGBoost) method with the Binary Relevance approach. To improve the model's performance, hyperparameter tuning was conducted using GridSearchCV. The dataset consisted of 7,000 translated Sahih Bukhari hadiths labeled into three categories: recommendation, prohibition, and information. The results show that the model achieved an accuracy of up to 94.64%, precision of 95.17%, recall of 99.47%, and F1-score of 97.24%. Furthermore, the implementation of hyperparameter tuning using GridSearchCV improved the recall from 32.72% to 35.66% and the F1-score from 47.09% to 50.00%, while reducing the hamming loss from 9.19% to 9.10%. These findings indicate that the combination of XGBoost, Binary Relevance, and GridSearchCV is effective for multi-label classification of translated Sahih Bukhari hadiths.

Keywords: Hadith; Multi-Label Classification; XGBoost; Binary Relevance; GridSearchCV

1. PENDAHULUAN

Dalam ajaran Islam, Al-Qur'an dan Hadis menjadi landasan utama yang digunakan sebagai pedoman dalam kehidupan sehari-hari. Hadis yang memuat perkataan serta perbuatan Nabi Muhammad SAW berperan sebagai pelengkap Al-Qur'an sekaligus memberikan penjelasan yang lebih rinci mengenai ajaran Islam [1]. Hadis perlu dipelajari dan diamalkan dalam kehidupan sehari-hari. Banyak ulama hadis yang meriwayatkan hadis, salah satunya adalah Imam Bukhari yang dikenal memiliki tingkat keabsahan hadis yang tinggi. Hadis-hadis Shahih riwayat Imam Bukhari memuat ajaran berupa anjuran yang perlu dilaksanakan, larangan yang harus dihindari, serta berbagai informasi yang penting bagi umat muslim [2].

Salah satu permasalahan dalam mempelajari hadis adalah menentukan jenis ajaran yang terkandung di dalamnya. Matan hadis tidak jarang memuat unsur anjuran, larangan, serta informasi lain secara bersamaan, sehingga hal ini dapat menyulitkan umat Islam dalam memahami dan mengamalkan kandungan hadis secara tepat. Pendekatan klasifikasi teks konvensional yang mengasumsikan satu label untuk setiap dokumen menjadi kurang sesuai, karena satu hadis dapat mencakup lebih dari satu kategori ajaran [3]. Permasalahan tersebut dapat diatasi melalui penerapan klasifikasi teks Multi-label. Klasifikasi Multi-label merupakan metode klasifikasi yang memungkinkan satu data teks memiliki lebih dari satu label kategori secara bersamaan. Pendekatan ini sesuai dengan karakteristik teks hadis yang dapat mengandung unsur anjuran, larangan, dan informasi [1].

Klasifikasi merupakan proses pengelompokan atau pemberian label terhadap data atau objek baru berdasarkan karakteristik tertentu [4]. Dalam pengembangannya, berbagai algoritma pembelajaran mesin telah digunakan untuk menyelesaikan permasalahan klasifikasi teks. salah satu algoritma yang dapat digunakan dalam proses klasifikasi adalah *Extreme Gradient Boosting* (XGBoost). Algoritma ini merupakan pengembangan dari metode *gradient boosting* yang dirancang untuk menghasilkan proses pelatihan model dengan waktu komputasi yang lebih cepat. Selain itu, XGBoost juga memiliki fitur regularisasi yang membantu mengurangi *overfitting* pada model [5]. Algoritma ini juga memiliki keunggulan dalam menangani data berskala kecil hingga besar serta mampu menyelesaikan berbagai tugas pembelajaran mesin, seperti klasifikasi, regresi, dan peringkat. Selain itu, XGBoost dikenal memiliki tingkat fleksibilitas yang baik dan mampu menghasilkan performa yang lebih tinggi dibandingkan algoritma *tea learning* konvensional [6].

Bulletin of Data Science
ISSN 2745-7966 (Media Online)
Vol 5, No 3, June 2026 | Page 170−178
DOI: 10.47065/bulletinds.v5i3.10142
https://ejurnal.seminar-id.com/index.php/bulletinds

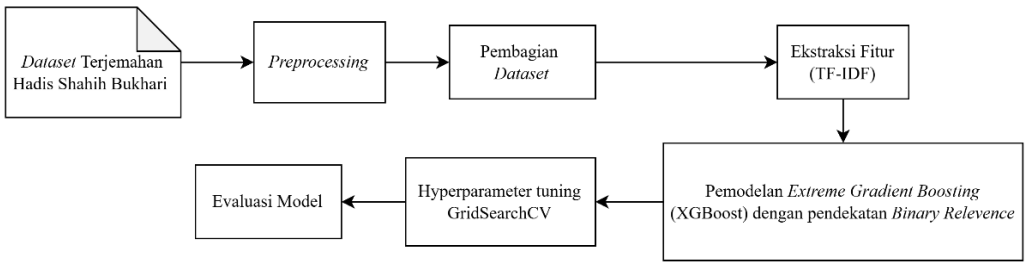
Penerapan algoritma XGBoost pada berbagai penelitian klasifikasi telah menunjukkan hasil yang cukup baik. Putri et al. [7] menerapkan algoritma XGBoost pada klasifikasi spam email dan memperoleh nilai akurasi sebesar 95,3%, precision 95,1%, recall 95,6%, dan F1-score 95,2%. Selanjutnya, Widiarta et al. [8] menggunakan XGBoost pada klasifikasi sentimen Twitter dengan data hasil augmentasi dan stemming dan memperoleh akurasi sebesar 85,27%. Hendrawan [9] menunjukkan kombinasi XGBoost dan Word2Vec menghasilkan F1-score sebesar 94,1%. Sementara itu, Fernando [10] menerapkan XGBoost pada klasifikasi cyberbullying di Twitter dan memperoleh akurasi sebesar 83,24%, serta menunjukkan performa yang lebih baik dibandingkan Support Vector Machine (SVM).Selain itu, Rusdi et al. [11] mengombinasikan XGBoost dengan Synthetic Minority Over-sampling Technique (SMOTE) dan menghasilkan nilai akurasi, precision, recall, serta F1-score di atas 98,88%.

Penelitian terkait klasifikasi Multi-label hadis terjemahan Shahih Bukhari juga telah dilakukan oleh beberapa peneliti sebelumnya. Ahmad et al. [1] membandingkan metode Random Forest dan Long Short-Term Memory (LSTM), diperoleh hasil bahwa metode random forest memberikan performa lebih baik. Selanjutnya, Kustiawan et al. [2] menerapkan metode CART dan esemble learning bagging menghasilkan nilai akurasi sebesar 80,86%. Selain itu, Ramadhani et al. [3] membandingkan metode Support Vector Machine (SVM) dan LSTM pada klasifikasi Multi-label hadis terjemahan Shahih Bukhari. Hasil penelitian tersebut menunjukkan bahwa LSTM memberikan performa yang lebih baik pada kategori anjuran dan larangan, sedangkan SVM lebih unggul pada kategori informasi.

Berdasarkan penelitian terdahulu tersebut, dapat diketahui bahwa berbagai metode *machine learning* dan *deep learning* telah diterapkan pada klasifikasi Multi-label hadis terjemahan Shahih Bukhari dan menunjukkan performa yang cukup baik. Namun, penerapan algoritma XGBoost pada klasifikasi Multi-label hadis terjemahan Shahih Bukhari masih belum banyak diteliti. Selain itu, penelitian sebelumnya lebih banyak berfokus pada perbandingan performa antar algoritma klasifikasi, sedangkan pengaruh optimasi *hyperparameter* terhadap performa model pada klasifikasi Multi-label hadis belum banyak dibahas. Penelitian ini bertujuan untuk mengklasifikasikan hadis terjemahan Shahih Bukhari ke dalam kategori anjuran, larangan, dan informasi menggunakan algoritma XGBoost dengan pendekatan *Binary Relevance*. *Dataset* yang digunakan terdiri dari 7000 data hadis berlabel. Selain itu, *hyperparameter tuning* menggunakan *GridSearchCV* diterapkan untuk memperoleh kombinasi parameter yang optimal serta mengevaluasi pengaruhnya terhadap performa model. Perbedaan penelitian ini dengan penelitian sebelumnya terletak pada penerapan algoritma XGBoost pada kasus klasifikasi Multi-label hadis terjemahan Shahih Bukhari serta analisis pengaruh *hyperparameter tuning* terhadap peningkatan performa model. Hasil penelitian ini diharapkan dapat memberikan alternatif metode yang efektif untuk klasifikasi Multi-label hadis serta menjadi referensi bagi penelitian selanjutnya yang menerapkan algoritma XGBoost pada pengolahan teks keagamaan.

2. METODOLOGI PENELITIAN

Tahapan penelitian yang dilakukan dalam penelitian ini dapat dilihat pada Gambar 1. Proses penelitian dimulai dari pengumpulan *dataset* hadis terjemahan Shahih Bukhari, kemudian dilakukan tahap *preprocessing* untuk mempersiapkan data sebelum digunakan dalam proses klasifikasi. Selanjutnya, *dataset* dibagi menjadi data latih dan data uji, kemudian dilakukan ekstraksi fitur menggunakan TF-IDF. Tahap berikutnya adalah pemodelan menggunakan algoritma *Extreme Gradient Boosting* (XGBoost) dengan pendekatan *Binary Relevance* serta optimasi parameter menggunakan *GridSearchCV*. Setelah model terbentuk, dilakukan evaluasi untuk mengukur performa klasifikasi yang dihasilkan.



Gambar 1. Metodologi Penelitian

2.1 *Dataset* Hadis Shahih Bukhari

Penelitian ini menggunakan dataset hadis Shahih Bukari terjemahan Bahasa Indonesia yang telah diberi label pada penelitian [1] berdasarkan kandungan makna yang terdapat pada setiap hadis. Proses pelabelan dilakukan dengan meninjau isi matan hadis untuk menentukan kategori anjuran, larangan, dan informasi. Tabel 1 merupakan representasi isi hadis yang digunakan dalam penelitian berdasarkan kategori anjuran, larangan, dan informasi.

Bulletin of Data Science
ISSN 2745-7966 (Media Online)
Vol 5, No 3, June 2026 | Page 170−178
DOI: 10.47065/bulletinds.v5i3.10142
https://ejurnal.seminar-id.com/index.php/bulletinds

Tabel 1. *Representasi Dataset Terjemahan Hadis Shahih Bukhari*

No.	Terjemah	Anjuran	Larangan	Informasi
45	Telah menceritakan kepada kami [Musaddad] berkata, telah menceritakan kepada kami [Yahya] dari [Isma'il] berkata, telah menceritakan kepadaku [Qais bin Abu Hazim] dari [Jarir bin Abdullah] berkata: "Aku telah membai'at Rasulullah untuk menegakkan shalat, menunaikan zakat dan menasehati kepada setiap muslim".	1	0	1
100	Dan Ali berkata, "Berbicaralah dengan manusia sesuai dengan kadar pemahaman mereka, apakah kalian ingin jika Allah dan rasul-Nya didustakan?" Telah menceritakan kepada kami ['Ubaidullah bin Musa] dari [Ma 'ruf bin Kharrabudz] dari [Abu Ath Thufail] dari ['Ali] seperti itu."	1	0	1
6931	Telah menceritakan kepada kami [Qutaibah bin Sa'id] dari [Malik] dari [Nafi'] dari [Ibnu Umar], Rasulullah Shallallahu'alaihiwasallam melarang (jual beli) najasy (penipuan).	0	1	0

Pada Tabel 1 ditampilkan contoh data hadis Terjemahan Hadis Shahih Bukhari yang telah dikategorikan ke dalam label anjuran, larangan, dan informasi. Nilai 1 menunjukkan bahwa teks hadis mengandung kategori terkait, sedangkan nilai 0 menunjukkan bahwa kategori tersebut tidak terkandung dalam teks hadis. Distribusi data pada masing-masing kategori disajikan pada tabel 2.

Tabel 2. Distribusi Data Hadis Shahih Bukhari

Kategori	Kelas	
	1	0
Anjuran	1.335	5.665
Larangan	844	6.156
Informasi	6.634	366

Berdasarkan Tabel 2, jumlah data pada setiap kategori menunjukkan distribusi yang tidak merata. Kategori informasi memiliki jumlah data paling banyak, sedangkan kategori anjuran dan larangan memiliki jumlah data yang lebih sedikit. Perbedaan distribusi tersebut mengindikasikan bahwa *dataset* yang digunakan tidak memiliki sebaran data yang merata pada setiap kategori.

2.2 *Text Preprocessing*

Text preprocessing merupakan proses pembersihan teks sebelum data digunakan pada tahap klasifikasi. Pada tahap ini, teks diolah menjadi bentuk yang lebih sederhana agar proses ekstraksi fitur dapat dilakukan dengan lebih baik [12]. Tahapan preprocessing teks dijelaskan sebagai berikut:

a. *Cleaning*: tahapan ini dilakukan dengan membersihkan teks dari simbol, angka, dan karakter yang tidak digunakan dalam proses pengolahan data.
b. *Case folding*: pada tahap ini semua teks diubah menjadi huruf kecil untuk menghindari perbedaan penulisan huruf kapital.
c. *Stopward Removal*: tahapan ini dilakukan dengan menghapus kata-kata yang sering muncul tetapi tidak memiliki pengaruh besar dalam proses klasifikasi.
d. *Tokenizing*: pada tahap ini teks dibagi menjadi bagian-bagian kata atau frasa.

Penelitian ini tidak menggunakan tahapan *stemming*, *sequence*, dan *padding* karena metode yang digunakan adalah XGBoost dengan pembobotan TF-IDF. *Sequence* dan *padding* umumnya digunakan pada model *deep learning* seperti LSTM yang memerlukan input berupa urutan angka dengan panjang seragam. Sementara itu, TF-DF secara langsung mengubah teks menjadi representasi numerik yang dapat diproses oleh XGBoost. Adapun *stemming* tidak diterapkan karena kata-kata hasil tokenisasi dipertahankan dalam bentuk aslinya agar informasi pada teks tetap terjaga selama proses klasifikasi.

2.3 Pembagian *Dataset*

Dataset pada penelitian ini dibagi menjadi data *training* dan data *testing* dengan perbandingan 80:20. Pembagian dilakukan untuk memperoleh data yang digunakan pada proses pelatihan model dan data yang digunakan pada tahap pengujian. Dari total 7.000 data hadis, sebanyak 5.600 data digunakan sebagai data *training*, sedangkan 1.400 data digunakan sebagai data *testing*. Data *training* digunakan untuk membangun model klasifikasi, sementara data testing digunakan untuk melihat hasil kinerja model terhadap data yang belum digunakan sebelumnya.

2.4 Esktraksi Fitur

Bulletin of Data Science
ISSN 2745-7966 (Media Online)
Vol 5, No 3, June 2026 | Page 170−178
DOI: 10.47065/bulletinds.v5i3.10142
https://ejurnal.seminar-id.com/index.php/bulletinds

Pada tahap ekstraksi fitur digunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*). Metode ini berfungsi untuk memberikan bobot pada setiap kata berdasarkan tingkat kemunculannya dalam dokumen. Kata yang sering muncul pada suatu dokumen tetapi jarang ditemukan pada dokumen lainnya akan memperoleh bobot yang lebih besar. TF-IDF digunakan untuk merepresentasikan teks ke dalam bentuk numerik sehingga dapat membantu proses identifikasi kata-kata yang dianggap penting dalam klasifikasi teks [13]. Rumus perhitungan TF-IDF ditunjukkan pada Persamaan 1-3.

$$\boldsymbol{TF_{(i,j)} = \frac{jumlah\ kemunculan\ kata\ i\ dalam\ dokumen\ j}{total\ kata\ dalam\ dokumen\ j}} \quad (1)$$

$$\boldsymbol{IDF_{(i,j)} = \log \frac{jumlah\ dokumen}{jumlah\ dokumen\ yang\ mengandung\ kata\ i}} \quad (2)$$

$$\boldsymbol{TF - IDF = TF \times IDF} \quad (3)$$

Term Frequency (TF) menunjukkan frekuensi kemunculan suatu kata dalam dokumen. Sementara itu, *Inverse Document Frequency* (IDF) digunakan untuk melihat tingkat penyebaran kata pada seluruh dokumen dalam *dataset*. Nilai TF-IDF diperoleh dari kombinasi nilai TF dan IDF yang digunakan untuk memberikan bobot pada setiap kata. Semakin besar nilai TF-IDF yang dihasilkan, semakin tinggi tingkat kepentingan kata tersebut dalam membedakan isi suatu dokumen pada proses klasifikasi.

2.5 *Extreme Gradient Boosting* (XGBoost)

Extreme Gradient Boosting (XGBoost) termasuk algoritma *machine laerning* berbasis *esemble* yang digunakan untuk menangani permasalahan klasifikasi dan regresi [13]. Pada tahun 2011, Tianqi Chen dan Carlos memperkenalkan algoritma XGBoost, kemudian terus dikembangkan dan disempurnakan oleh berbagai peneliti dalam studi-studi selanjutnya [14]. XGBoost menggunakan model yang lebih terstruktur dalam proses pembentukan pohon regresi sehingga dapat meningkatkan kinerja model serta membantu mengurangi kompleksitas yang dapat menyebabkan *overfitting*. Hasil prediksi akhir diperoleh dari penjumlahan seluruh prediksi pohon regresi yang terbentuk [5]. *Extreme Gradient Boosting* (XGBoost) juga termasuk salah satu varian *boosting* yang dapat berjalan 10 kali lebih cepat daripada implementasi *Gradient Boosting* lainnya [15]. Pada penelitian ini, metode XGBoost diterapkan menggunakan pendekatan *Binary Relevence* untuk menangani klasifikasi Multi-label. *Binary Relevence* dipilih karena mampu membangun proses klasifikasi secara terpisah untuk setiap label pada klasifikasi Multi-label. Pendekatan ini relatif sederhana untuk diterapkan serta memiliki tingkat kompleksitas yang meningkat secara linear sesuai dengan jumlah label yang digunakan. Selain itu, *Binary Relevence* dapat dikombinasikan dengan berbagai metode pembelajaran mesin untuk menangani masing-masing label secara individual [16].

2.6 *Hyperparameter Tuning* dengan *Grid Search Cross Validation*

Hyperparameter tuning berperan penting dalam meningkatkan performa algoritma pembelajaran mesin[17]. *hyperparameter tuning* dilakukan untuk memperoleh kombinasi *hyperparameter* terbaik untuk meningkatkan performa model [18]. Salah satu metode *hyperparameter tuning* yang digunakan pada penelitian ini adalah *Grid Search Cross Validation* (GridSearchCV). Metode ini bekerja dengan mencoba beberapa kombinasi *hyperparameter* untuk melihat parameter yang memberikan hasil evaluasi terbaik. Dalam prosesnya, *GridSearchCV* melakukan pelatihan model menggunakan kombinasi *hyperparameter* yang berbeda, kemudian mengevaluasinya menggunakan *Cross Validation.* Kombinasi *hyperparameter* dengan hasil evaluasi terbaik selanjutnya dipilih sebagai model [19]. Berdasarkan penelitian [5]. Penggunaan *hyperparameter tuning* menggunakan *GridSearchCV* pada algoritma XGBoost memberikan hasil evaluasi yang lebih baik dibandingkan parameter *default*. Peningkatan nilai *accuracy*, *precision*, dan *recall* menunjukkan bahwa proses *tuning* membantu meningkatkan performa klasifikasi model. Proses *tuning* pada penelitian ini dilakukan menggunakan beberapa parameter utama XGBoost, yaitu *n_estimator*, *max_depth*, dan *learning_rate*. Nilai yang diuji pada *n_estimator* adalah 100 dan 200 untuk mengatur jumlah pohon keputusan yang dibangun, *max_depth* diuji pada nilai 4 dan 5 untuk mengontrol kedalaman pohon, *leraning_rate* diuji pada nilai 0.01, 0.05, dan 0.1 untuk mengatur kecepatan pembelajaran model. Rentang nilai tersebut ditentukan dengan mempertimbangkan jumlah *dataset* yang digunakan sehingga proses pencarian kombinasi *hyperparameter* terbaik dapat dilakukan tanpa membutuhkan waktu yang terlalu lama. Kombinasi nilai yang diperoleh kemudian digunakan untuk menghasilkan performa model yang lebih optimal dalam proses klasifikasi Multi-label.

2.7 Evaluasi Model

Tahap evaluasi dilakukan untuk mengetahui performa model XGBoost dengan pendekatan *Binary Relevence* dalam klasifikasi hadis ke dalam kategori anjuran, larangan, dan informasi. Pengujian model dilakukan menggunakan data testing sebanyak 1.400 hadis atau 20% dari total *dataset* yang telah dipisahkan sebelumnya dari data *training*. Pada tahap ini digunakan beberapa metrik evaluasi, yaitu *accuracy*, *precision*, *recall* F1-score, dan *hamming loss*. Penggunaan beberapa metrik evaluasi tersebut diperlukan untuk memperoleh gambaran kinerja model yang lebih menyeluruh. Hal ini

Bulletin of Data Science
ISSN 2745-7966 (Media Online)
Vol 5, No 3, June 2026 | Page 170−178
DOI: 10.47065/bulletinds.v5i3.10142
https://ejurnal.seminar-id.com/index.php/bulletinds

karena nilai *accuracy* tidak selalu mampu merepresentasikan performa model secara lengkap, terutama pada data yang memiliki distribusi kelas tidak seimbang. Dalam kondisi tersebut, model dapat menghasilkan nilai *accuracy* yang tinggi, tetapi belum tentu mampu mengklasifikasikan setiap kelas dengan baik. Oleh sebab itu, metrik *precision*, *recall*, F1-score dan *hamming loss* digunakan untuk mengevaluasi tingkat ketepatan serta kemampuan model dalam mengenali data pada masing-masing label. Kombinasi metrik tersebut memungkinkan proses evaluasi dilakukan secara lebih komprehensif sehingga menghasilkan penilaian yang lebih objektif terhadap performa model klasifikasi [20].Tahap evaluasi juga digunakan untuk membandingkan performa model sebelum dan sesudah *hyperparameter tuning*. Perhitungan metrik evaluasi ditunjukkan pada persamaan berikut.

$$\boldsymbol{Accuracy = \frac{TP + TN}{TP + TN + FP + FN}} \quad (4)$$

Accuracy digunakan untuk mengukur tingkat ketepatan model secara keseluruhan dengan membandingkan jumlah prediksi yang benar, yaitu *True Positive* (TP) dan *True Negative* (TN), terhadap seluruh data pengujian yang terdiri dari *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN).

$$\boldsymbol{Precision = \frac{TP}{TP + FP}} \quad (5)$$

Precision digunakan untuk mengukur tingkat ketepatan model dalam memberikan prediksi positif dengan membandingkan jumlah *True Positive* (TP) terhadap seluruh data yang diprediksi sebagai positif, yaitu *True Positive* (TP) dan *False Positive* (FP).

$$\boldsymbol{Recall = \frac{TP}{TP + FN}} \quad (6)$$

Recall digunakan untuk mengukur kemampuan model dalam mengidentifikasi seluruh data yang benar-benar termasuk dalam kelas positif dengan membandingkan jumlah *True Positive* (TP) terhadap seluruh data positif aktual, yaitu *True Positive* (TP) dan *False Negative* (FN).

$$\boldsymbol{F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}} \quad (7)$$

F1-score merupakan ukuran evaluasi yang menggabungkan precision dan recall menjadi satu nilai tunggal, sehingga memberikan keseimbangan antara ketepatan prediksi dan kemampuan model dalam mendeteksi kelas positif.

$$\boldsymbol{Hamming\ Loss = \frac{Jumlah\ label\ yang\ salah}{Jumlah\ data \times Jumlah\ kategori}} \quad (8)$$

Hamming Loss digunakan untuk mengukur tingkat kesalahan pada klasifikasi Multi-label dengan membandingkan jumlah label yang salah diprediksi terhadap total data pengujian dan jumlah kategori label yang ada. Semakin kecil nilai *Hamming Loss*, maka semakin baik kinerja model dalam melakukan klasifikasi.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengujian dan Evaluasi Model

Pada penelitian ini, proses pengolahan dan klasifikasi data hadis terjemahan Shahih Bukhari dilakukan menggunakan bahasa pemrograman *Python*. *Dataset* yang digunakan berupa 7.000 data hadis yang telah dilabeli ke dalam tiga kategori, yaitu anjuran, larangan, dan informasi. Setiap data direpresentasikan dalam bentuk Multi-label, di mana satu data dapat memiliki lebih dari satu label secara bersamaan. Tahapan awal dalam penelitian ini adalah *preprocessing* data yang bertujuan untuk membersihkan dan mempersiapkan data teks sebelum digunakan pada proses klasifikasi. *Preprocessing* dilakukan melalui beberapa tahap, yaitu, *cleaning, case folding, stopward removal*, serta *tokenizing*. Hasil dari proses preprocessing tersebut ditunjukkan pada Tabel 3 berikut.

Tabel 3. Hasil *Preprocessing*

No	Hadis	Anjuran	Larangan	Informasi	Hadis Bersih
45	Telah menceritakan kepada kami [Musaddad] berkata, telah menceritakan kepada kami [Yahya] dari [Isma'il] berkata, telah menceritakan kepadaku [Qais bin Abu Hazim] dari [Jarir bin Abdullah] berkata: "Aku telah membai'at Rasulullah untuk menegakkan shalat, menunaikan zakat dan menasehati kepada setiap muslim".	1	0	1	kepadaku aku membai at rasulullah menegakkan shalat menunaikan zakat menasehati muslim

Bulletin of Data Science
ISSN 2745-7966 (Media Online)
Vol 5, No 3, June 2026 | Page 170−178
DOI: 10.47065/bulletinds.v5i3.10142
https://ejurnal.seminar-id.com/index.php/bulletinds

No	Hadis	Anjuran	Larangan	Informasi	Hadis Bersih
100	Dan Ali berkata, "Berbicaralah dengan manusia sesuai dengan kadar pemahaman mereka, apakah kalian ingin jika Allah dan rasul-Nya didustakan?" Telah menceritakan kepada kami ['Ubaidullah bin Musa] dari [Ma 'ruf bin Kharrabudz] dari [Abu Ath Thufail] dari ['Ali] seperti itu."	1	0	1	ali berbicaralah manusia sesuai kadar pemahaman kalian allah rasul nya didustakan
6931	Telah bercerita kepada kami [Musaddad] dari [Yahya] dari ['Ubaidullah] berkata, telah bercerita kepadaku [Nafi'] dari [Ibnu 'Umar radliallahu 'anhuma] dari Nabi shallallahu 'alaihi wasallam bersabda: "Penyakit panas (demam) berasal dari didihan api jahannam maka redakanlah dengan air".	1	0	1	bercerita bercerita kepadaku nabi shallallahu alaihi wasallam bersabda penyakit panas demam berasal didihan api jahannam redakanlah air

Hasil *preprocessing* yang ditampilkan pada Tabel 3 menunjukkan bahwa data teks telah dibersihkan dari unsur-unsur yang dapat menimbulkan *noise* pada proses klasifikasi. Dengan demikian, informasi yang tersisa lebih berfokus pada kandungan makna hadis sehingga dapat membantu model dalam mengenali pola yang berkaitan dengan kategori anjuran, larangan, dan informasi. Setelah tahap *preprocessing* selesai, data kemudian digunakan sebagai *input* untuk proses klasifikasi. Penelitian ini menggunakan pendekatan klasifikasi Multi-label untuk mengelompokkan hadis ke dalam tiga kategori, yaitu anjuran, larangan, dan informasi. Dalam penerapannya, metode Multi-label yang digunakan adalah *Binary Relevance*, yaitu setiap label diperlakukan sebagai masalah klasifikasi biner yang terpisah. Dengan pendekatan ini, model dibangun secara independen untuk masing-masing kategori sehingga setiap label memiliki prediksi sendiri tanpa saling bergantung.

Algoritma yang digunakan dalam proses klasifikasi adalah *Extreme Gradient Boosting* (XGBoost), yaitu algoritma berbasis *ensemble learning* yang menggabungkan beberapa model *decision tree* untuk meningkatkan performa prediksi. Dalam penelitian ini, XGBoost digunakan untuk mempelajari pola dari data yang telah melalui proses *preprocessing* dan diubah menjadi representasi fitur. Model kemudian dilatih menggunakan data *training* sebesar 80% dan data testing sebesar 20% dari total *dataset* untuk mengevaluasi kinerja model. Pengujian dilakukan dalam dua skenario, yaitu menggunakan model XGBoost tanpa *hyperparameter tuning* dan menggunakan *hyperparameter tuning* dengan metode *GridSearchCV*. Pada skenario pertama, model menggunakan parameter bawaan dari *library* XGBoost tanpa dilakukan proses optimasi parameter. Seluruh data *training* digunakan untuk membangun model, sedangkan data testing digunakan untuk mengukur performa klasifikasi pada masing-masing kategori label. Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan F1-score. Hasil evaluasi model XGBoost tanpa *hyperparameter tuning* disajikan pada Tabel 4.

Tabel 4. Hasil Evaluasi XGBoost Tanpa *HyperparameterTuning*

Kategori	Accuracy	Precision	Recall	F-1 score
Anjuran	85,71%	83,96%	32,72%	47,09%
Larangan	92,07%	78,95%	50,85%	61,86%
Informasi	94,64%	95,11%	99,47%	97,24%

Berdasarkan Tabel 4, kategori informasi memperoleh performa terbaik dengan nilai *accuracy* sebesar 94,64%, *precision* sebesar 95,11%, *recall* sebesar 99,47%, dan F1-score sebesar 97,24%. Sementara itu, kategori larangan memperoleh F1-score sebesar 61,86%, sedangkan kategori anjuran memperoleh F1-score terendah sebesar 47,09% dengan nilai *recall* sebesar 32,72%. Hasil pada Tabel 4 menunjukkan bahwa model XGBoost tanpa *hyperparameter tuning* lebih mampu mengenali kategori informasi dibandingkan kategori anjuran dan larangan.

Pada skenario kedua, dilakukan proses *hyperparameter tuning* menggunakan metode *GridSearchCV* untuk memperoleh kombinasi parameter terbaik. Parameter yang diuji meliputi *n_estimator* = (100, 200), *max_depth* = (4,5), dan *learning rate* = (0.01, 0.05, 0.1). Setiap kombinasi parameter diuji untuk menentukan konfigurasi yang memberikan performa terbaik pada model. Hasil kombinasi hyperparameter terbaik dapat dilihat pada tabel 5 berikut.

Tabel 5. Hasil Kombinasi Hyperparameter Terbaik

Hyperparameter	GridSearchValues	Nilai Hyperparameter Terbaik
n_estimator	100, 200	200
max_depth	4, 5	5
learning_rate	0.01, 0.05, 0.1	0.1

Berdasarkan Tabel 5, proses *GridSearchCV* menghasilkan kombinasi *hyperparameter* terbaik, yaitu *n_estimators* = 200, *max_depth* = 5, dan *learning_rate* = 0,1. Kombinasi parameter tersebut selanjutnya digunakan pada

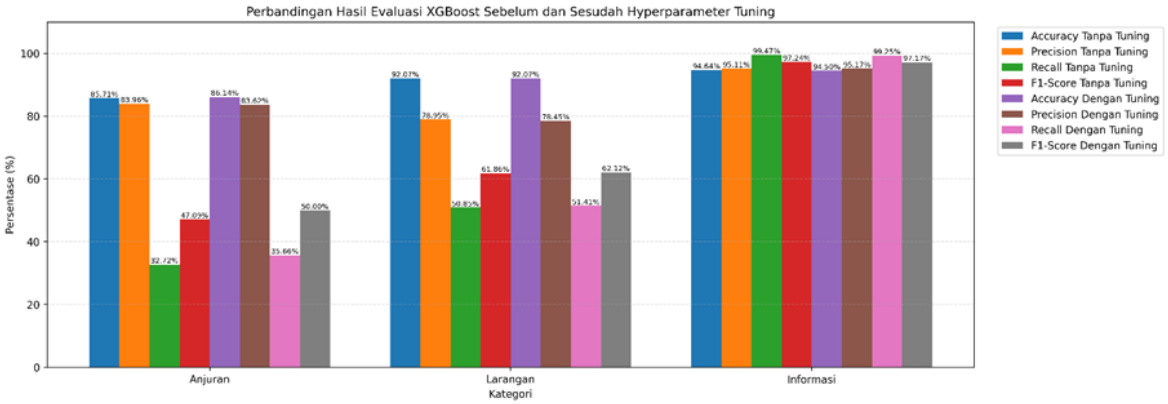
Bulletin of Data Science
ISSN 2745-7966 (Media Online)
Vol 5, No 3, June 2026 | Page 170−178
DOI: 10.47065/bulletinds.v5i3.10142
https://ejurnal.seminar-id.com/index.php/bulletinds

proses pelatihan dan pengujian model XGBoost dengan menggunakan data *training* dan testing yang sama seperti pada skenario sebelumnya. Hasil evaluasi model setelah penerapan *hyperparameter tuning* ditunjukkan pada Tabel 6.

Tabel 6. Hasil Evaluasi XGBoost Dengan *Hyperparameter Tuning*

Kategori	Accuracy	Precision	Recall	F1-score
Anjuran	86,14%	83,62%	35,66%	50,00%
Larangan	92,07%	78,45%	51,41%	62,12%
Informasi	94,50%	95,17%	99,25%	97,17%

Berdasarkan hasil evaluasi pada Tabel 6, penerapan *hyperparameter tuning* memberikan perubahan pada beberapa kategori. Pada kategori anjuran, nilai *recall* meningkat dari 32,72% menjadi 35,66%, dan nilai F1-score meningkat dari 47,09% menjadi 50,00%. Peningkatan juga terjadi pada kategori larangan dengan nilai *recall* sebesar 51,41% dan F1-score sebesar 62,12%. Hal ini menunjukkan bahwa proses *tuning* membantu model mengenali data pada kedua kategori tersebut dengan lebih baik dibandingkan sebelumnya. Pada kategori informasi, model tetap memperoleh nilai evaluasi tertinggi dengan *accuracy* sebesar 94,50%, *precision* 95,17%, *recall* 99,25% dan F1-score 97,17%. Meskipun beberapa nilai evaluasi mengalami sedikit penurunan setelah *tuning*, performa pada kategori informasi masih tergolong sanga baik karena seluruh nilai evaluasi berada di atas 90%. Perbandingan hasil evaluasi model sebelum dan sesudah *hyperparameter tuning* ditunjukkan pada gambar 2.



Gambar 2. Perbandingan Hasil Evaluasi XGBoost Sebelum dan Sesudah *Hyperparameter Tuning*

Berdasarkan Gambar 2, perubahan performa paling terlihat pada kategori anjuran dan larangan, terutama pada nilai *recall* dan F1-score. Kondisi tersebut Menunjukkan bahwa optimasi parameter membantu meningkatkan kemampuan model dalam mengenali data pada kedua kategori tersebut. Nilai *accuracy* pada seluruh kategori cenderung stabil sehingga proses *tuning* tidak memberikan perubahan yang terlalu besar terhadap keseluruhan performa model secara keseluruhan.

Selain menggunakan *accuracy*, *precision*, *recall*, dan F1-score, evaluasi model juga dilakukan menggunakan *Hamming Loss* untuk mengukur tingkat kesalahan prediksi pada setiap label. Hasil perhitungan *Hamming Loss* sebelum penerapan *hyperparameter tuning* dapat dilihat pada Tabel 7.

Tabel 7. Hasil *Hamming Loss* Sebelum *Hyperparameter Tuning*

Kategori	Benar	Salah	Total
Anjuran	1200	200	1400
Larangan	1289	111	1400
Informasi	1325	75	1400
Total	3814	386	4200
Hamming Loss	0,091904762		9,19%

Hasil perhitungan *Hamming Loss* setelah penerapan *hyperparameter tuning* ditunjukkan pada Tabel 8.

Tabel 8. Hasil *Hamming Loss* Setelah *Hypperparameter Tuning*

Kategori	Benar	Salah	Total
Anjuran	1206	194	1400
Larangan	1289	111	1400
Informasi	1323	77	1400
Total	3818	382	4200
Hamming Loss	0,090952381		9,10%

Tabel 7 dan Tabel 8, menunjukkan hasil nilai *hamming loss* pada model XGBoost sebelum dan sesudah penerapan *hyperparameter tuning* menggunakan *GridSearchCV*. Nilai *hamming loss* digunakan untuk melihat tingkat

Bulletin of Data Science
ISSN 2745-7966 (Media Online)
Vol 5, No 3, June 2026 | Page 170−178
DOI: 10.47065/bulletinds.v5i3.10142
https://ejurnal.seminar-id.com/index.php/bulletinds

kesalahan prediksi model pada klasifikasi Multi-label. Semakin kecil nilai *hamming loss* yang diperoleh, maka kemampuan model dalam melakukan klasifikasi pada setiap kategori label akan semakin baik.

Pada pemodelan tanpa *hyperparameter tuning*, model menghasilkan total prediksi benar sebanyak 3814 data dan prediksi salah sebanyak 386 data dari total 4200 label pengujian. Hasil tersebut menghasilkan nilai *hamming loss* sebesar 0,091904762 atau 9,19%. Kesalahan prediksi terbesar terdapat pada kategori anjuran dengan jumlah prediksi salah sebanyak 200 data. Sedangkan, pada pemodelan dengan *hyperparameter tuning* jumlah prediksi benar meningkat menjadi 3818 data dan jumlah prediksi salah menurun menjadi 382 data. Nilai *hamming loss* yang diperoleh juga mengalami penurunan menjadi 0,090952381 atau 9,10%. Peningkatan jumlah prediksi benar terlihat pada kategori anjuran yang meningkat dari 1200 data menjadi 1206 data setelah proses *tuning* dilakukan.

Hasil tersebut menunjukkan bahwa penerapan *hyperparameter tuning* menggunakan *GridSearchCV* memberikan pengaruh terhadap peningkatan performa model XGBoost dalam klasifikasi Multi-label hadis terjemahan Shahih Bukhari. Meskipun peningkatan yang diperoleh tidak terlalu signifikan, proses *tuning* mampu mengurangi tingkat kesalahan prediksi model sehingga menghasilkan performa klasifikasi yang lebih baik.

3.2 Pembahasan

Hasil penelitian menunjukkan bahwa algoritma XGBoost dengan pendekatan *Binary Relevence* mampu digunakan untuk melakukan klasifikasi Multi-label pada hadis terjemahan Shahih Bukhari ke dalam kategori anjuran, larangan, dan informasi. Berdasarkan hasil pengujian, kategori informasi memperoleh performa terbaik dibandingkan kategori lainnya. Sebaliknya, kategori anjuran dan larangan menunjukkan nilai *recall* yang lebih rendah. Kondisi tersebut dipengaruhi oleh distribusi data yang tidak seimbang, di mana jumlah data pada kategori informasi jauh lebih banyak dibandingkan kategori anjuran dan larangan. Akibatnya, model cenderung lebih mudah mengenali pola pada kategori informasi. Jika dibandingkan dengan penelitian klasifikasi Multi-label hadis terjemahan Shahih Bukhari yang dilakukan oleh Ahmad et al. [1], Kustiawan et al. [2], dan Ramadhani et al. [3] penelitian ini memberikan kesimpulan yang serupa bahwa metode *machine learning* dapat dimanfaatkan untuk mengelompokkan hadis ke dalam lebih dari satu kategori. Perbedaannya terletak pada algoritma yang digunakan. Penelitian terdahulu memanfaatkan Random Forest, CART, Bagging, SVM, dan LSTM, sedangkan penelitian ini menggunakan XGBoost dengan pendekatan *Binary Relevance*. Berdasarkan hasil evaluasi yang diperoleh, XGBoost mampu memberikan performa yang baik, terutama pada kategori informasi yang memiliki nilai evaluasi tertinggi dibandingkan kategori lainnya.

Penerapan *hyperparameter tuning* menggunakan *GridSearchCV j*uga memberikan pengaruh terhadap performa model. Setelah proses *tuning* dilakukan, nilai *recall* pada kategori anjuran meningkat dari 32,72% menjadi 35,66%, sedangkan pada kategori larangan meningkat dari 50,85% menjadi 51,41%. Selain itu, nilai *Hamming Loss* menurun dari 9,19% menjadi 9,10%. Kondisi tersebut menunjukkan bahwa optimasi parameter membantu model dalam mengurangi kesalahan prediksi dan meningkatkan kemampuan klasifikasi. Peningkatan performa setelah proses *tuning* ini juga ditemukan pada penelitian Herni Yulianti et al.[5] yang menyatakan bahwa penggunaan *GridSearchCV* pada algoritma XGBoost memberikan hasil yang lebih baik dibandingkan penggunaan parameter *default.* Meskipun peningkatan yang diperoleh pada penelitian ini tidak terlalu besar, proses *tuning* tetap memberikan kontribusi positif terhadap performa model yang dihasilkan.

4. KESIMPULAN

Penelitian ini berhasil menerapkan metode XGBoost dengan pendekatan *Binary Relevance* untuk melakukan klasifikasi Multi-label pada hadis terjemahan Shahih Bukhari ke dalam kategori anjuran, larangan, dan informasi. Hasil penelitian menunjukkan bahwa penerapan *hyperparameter tuning* menggunakan *GridSearchCV* mampu meningkatkan kinerja model dan mengurangi tingkat kesalahan prediksi dibandingkan penggunaan parameter bawaan. Dengan demikian, optimasi parameter memberikan kontribusi positif terhadap performa klasifikasi Multi-label menggunakan XGBoost. Meskipun demikian, penelitian ini masih memiliki keterbatasan karena hanya menggunakan algoritma XGBoost tanpa melakukan perbandingan dengan metode klasifikasi lainnya pada *dataset* yang sama. Selain itu, kemampuan model dalam mengenali seluruh data pada setiap kategori belum sepenuhnya optimal. Oleh karena itu, penelitian selanjutnya dapat mengkaji penggunaan algoritma lain maupun metode optimasi parameter yang berbeda untuk memperoleh performa klasifikasi yang lebih baik.

REFERENCES

[1] R. Z. N. Ahmad, N. S. Harahap, S. Agustian, I. Iskandar, and S. Sanjaya, "Perbandingan Performa Random Forest dan Long Short-Term Memory dalam Klasifikasi Teks Multilabel Terjemahan Hadits Bukhari," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 3, pp. 862–874, 2025, doi: 10.57152/malcom.v5i3.2046.

[2] R. Kustiawan, A. Adiwijaya, and M. D. Purbolaksono, "A Multi-label Classification on Topic of Hadith Verses in Indonesian Translation using CART and Bagging," *J. Media Inform. Budidarma*, vol. 6, no. 2, p. 868, 2022, doi: 10.30865/mib.v6i2.3787.

[3] A. Ramadhani, N. Safaat, S. Agustian, I. Iskandar, and S. Sanjaya, "Perbandingan Performa Metode Klasifikasi Teks Multilabel Hadis Terjemahan Bukhari Menggunakan Support Vector Machine dan Long Short Term Memory: Performance Comparison of

ISSN 2745-7966 (Media Online)
Vol 5, No 3, June 2026 | Page 170−178
DOI: 10.47065/bulletinds.v5i3.10142
https://ejurnal.seminar-id.com/index.php/bulletinds

Multilabel Text Classification Methods on Translated Hadiths of Bukhari Using Suppor," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 3 SE-, pp. 896–907, 2025, [Online]. Available: https://www.journal.irpi.or.id/index.php/malcom/article/view/2051

[4] A. Pebdika, R. Herdiana, and D. Solihudin, "Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima Pip," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 452–458, 2023, doi: 10.36040/jati.v7i1.6303.

[5] S. E. Herni Yulianti, Oni Soesanto, and Yuana Sukmawaty, "Penerapan Metode Extreme Gradient Boosting (XGBOOST) pada Klasifikasi Nasabah Kartu Kredit," *J. Math. Theory Appl.*, vol. 4, no. 1, pp. 21–26, 2022.

[6] O. Opitasari, F. Natsir, and E. S. Marsiani, "Klasifikasi Diagnosis untuk Penyakit Kanker Serviks Menggunakan Algoritma Extreme Gradient Boosting (XGBoost)," *J. Ilm. FIFO*, vol. 16, no. 1, p. 55, 2024, doi: 10.22441/fifo.2024.v16i1.006.

[7] L. G. A. Putri, S. A. Wicaksono, and B. Rahayudi, "Analisis Klasifikasi Spam Email Menggunakan Metode Extreme Gradient Boosting (XGBoost)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 2, pp. 1–8, 2025.

[8] I. Putu, A. Purnama Widiarta, R. Dwiyansaputra, and A. Aranta, "Analisis Sentimen Masyarakat Terhadap Kebijakan Penarapan PPKM Di Media Sosial Twitter Dengan Menggunakan Metode XGBOOST (Analysis Of Community Sentiment On The Policy Of Implementation Of PPKM On Twitter Social Media Using Xgboost Method)," *J. Teknol. Informasi, Komput. dan Apl.*, vol. 5, no. 2, pp. 154–163, 2023, [Online]. Available: http://jtika.if.unram.ac.id/index.php/JTIKA/

[9] I. R. Hendrawan, "Perbandingan Algoritma Naive Bayes, SVM Dan XGBOOST Dalam Klasifikasi Teks Sentimen Masyarakat Terhadap Produk Lokal Di Indonesia," *J. Transform.*, vol. 18, no. 1, pp. 1–8, 2022.

[10] Felix Fernando, "Klasifikasi Tweet Cyberbullying Dengan Menggunakan Algoritma Svm Dan Xgboost," *J. Ilmu Komput. dan Sist. Inf.*, vol. 13, no. 1, 2025, doi: 10.24912/jiksi.v13i1.32857.

[11] E. S. Rusdi *et al.*, "Comparison of Ann , Random Forest , and Xgboost in Antibiotic," vol. 12, no. 4, 2025, [Online]. Available: https://jtiik.ub.ac.id/index.php/jtiik/article/view/9487

[12] N. S. Putri, P. P. Adikara, and T. N. Fatyanosa, "Klasifikasi Emosi Multi Label Di Teks Bahasa Inggris Menggunakan Metode Naïve Bayes Dan Fitur Term Frequency," vol. 10, no. 1, pp. 1–9, 2026.

[13] E. Applications, "Implementation of TF-IDF and XGBoost Algorithms in Scientific Paper Classification," vol. 5, no. 1, pp. 1–5, 2025.

[14] M. Noorunnahar, A. H. Chowdhury, and F. A. Mila, "A tree based eXtreme Gradient Boosting (XGBoost) machine learning model to forecast the annual rice production in Bangladesh," *PLoS One*, vol. 18, no. 3 March, pp. 1–15, 2023, doi: 10.1371/journal.pone.0283452.

[15] M. R. Kurniawanda and F. A. T. Tobing, "Analysis Sentiment Cyberbullying In Instagram Comments with XGBoost Method," *IJNMT (International J. New Media Technol.*, vol. 9, no. 1, pp. 28–34, 2022, doi: 10.31937/ijnmt.v9i1.2670.

[16] E. B. Gulcan, I. S. Ecevit, and F. Can, "Binary Transformation Method for Multi-Label Stream Classification," *Int. Conf. Inf. Knowl. Manag. Proc.*, no. 3, pp. 3968–3972, 2022, doi: 10.1145/3511808.3557553.

[17] W. Nugraha and A. Sasongko, "Hyperparameter Tuning on Classification Algorithm with Grid Search," *Sistemasi*, vol. 11, no. 2, p. 391, 2022, doi: 10.32520/stmsi.v11i2.1750.

[18] I. Muhamad Malik Matin, "Hyperparameter Tuning Menggunakan GridsearchCV pada Random Forest untuk Deteksi Malware," *Multinetics*, vol. 9, no. 1, pp. 43–50, 2023, doi: 10.32722/multinetics.v9i1.5578.

[19] A. F. D. Putra, M. N. Azmi, H. Wijayanto, S. Utama, and I. G. P. W. Wedashwara Wirawan, "Optimizing Rain Prediction Model Using Random Forest and Grid Search Cross-Validation for Agriculture Sector," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 23, no. 3, pp. 519–530, 2024, doi: 10.30812/matrik.v23i3.3891.

[20] S. Wijaya and S. Aisa, "Model Klasifikasi Kenaikan Pangkat Pegawai Negeri Sipil Menggunakan," vol. 5, no. 1, pp. 1–9, 2025, doi: 10.47065/bulletinds.v5i1.9644.