InComTech: Jurnal Telekomunikasi dan Komputer
vol.15, no.2, Agustus 2025, 168-185
http://publikasi.mercubuana.ac.id/index.php/Incomtech
P-ISSN: 2085-4811 E-ISSN: 2579-6089

Perbandingan Klasifikasi Single-Label dan Multi-Label Ulasan Pengguna Lapangan Futsal di Semarang Menggunakan SVM

Achrijal Shohib Arya Syifa[1], Khotibul Umam[2*], Maya Rini Handayani[3], Siti Nur Aini[4]

[1,2,3,4] *Teknologi Informasi, Universitas Islam Negeri Walisongo, Jl. Prof. Dr. Hamka, Ngaliyan, Kota Semarang*
*Email Penulis Koresponden: khothibul_umam@walisongo.ac.id

***Abstrak*:**
Futsal merupakan cabang olahraga yang semakin populer di seluruh Indonesia, termasuk di Semarang. Penelitian ini bertujuan untuk melakukan klasifikasi sentimen ulasan pengguna mengenai lapangan futsal di Kota Semarang menggunakan metode *Support Vector Machine*. Data penelitian diperoleh melalui scraping ulasan *Google Maps* dengan ekstensi *Chrome Instant Data Scraper* dan terdiri dari 1.189 ulasan. Proses penelitian mencakup pengumpulan data, *cleaning* dan *pre-processing* (normalisasi teks, modifikasi data, tokenisasi, *stop word filtering*, *stemming*), pelabelan (*single-label* dan *multi-label*), pembagian data (80% pelatihan dan 20% pengujian), pemodelan menggunakan SVM (*single-label* dengan *GridSearchCV* dan *multi-label* dengan *One-vs-Rest Classifier*), serta evaluasi model dengan metrik presisi, *recall*, dan *F1-Score*. Hasil menunjukkan pemodelan *Support Vector Machine single-label* mencapai presisi 0.84, *recall* 0.73, dan *F1-Score* 0.78. Sementara pemodelan *Support Vector Machine multi-label* mencapai presisi 0.96, *recall* 0.88, dan *F1-Score* 0.92. Dari ulasan yang dianalisis, sebaran data pada *single-label* maupun *multi-label* menunjukan dominasi ulasan label fasilitas, menegaskan bahwa label fasilitas merupakan kategori yang paling sering dikomentari oleh pengguna. Temuan ini tidak hanya memberikan wawasan praktis bagi pengelola lapangan futsal, tetapi juga berkontribusi pada pengembangan metode klasifikasi ulasan berbasis machine learning dalam domain analisis opini, khususnya dalam membandingkan performa pendekatan single-label dan multi-label pada data multi-kategori di bidang teknologi informasi.

***Kata Kunci*:**
Klasifikasi Multi-Label, Single-Label, Ulasan Pengguna, Lapangan Futsal, Support Vector Machine

Riwayat Artikel:
Diserahkan 12 Juni 2025
Direvisi 06 Agustus 2025
Diterima 21 Agustus 2025

DOI:
10.22441/incomtech.v15i2.33905

1. PENDAHULUAN

Futsal merupakan cabang olahraga yang semakin populer di seluruh Indonesia, termasuk di Semarang. Semarang adalah salah satu kota dengan banyak penggemar futsal. Ini ditunjukkan oleh banyaknya orang yang terlibat dalam acara futsal, termasuk PORPROV (Pekan Olahraga antar Provinsi) Jawa Tengah 2023 [1]. Perkembangan olahraga ini mendorong pemilik fasilitas futsal meningkatkan sarana dan prasarana agar lebih diminati pengunjung. Oleh karena itu, penting untuk menganalisis opini pengguna terhadap lapangan futsal tersebut.

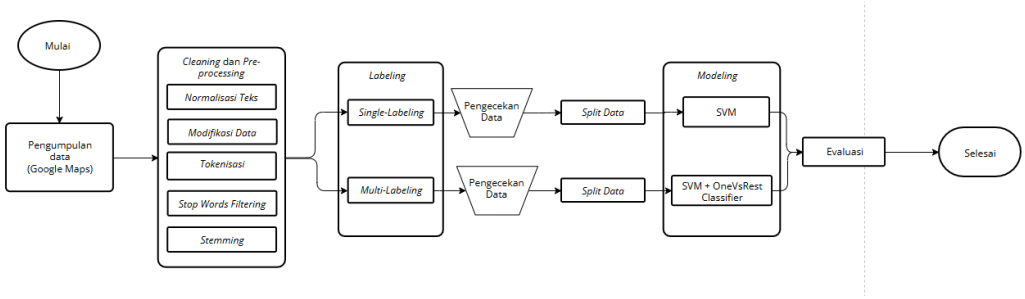
Di era digital, ulasan pengguna tentang fasilitas umum kini dapat diperoleh melalui platform online seperti *Google Maps*. *Google Maps* menyediakan detail lokasi, jarak, dan ulasan pengunjung untuk tempat-tempat yang dikunjungi [2]. Beberapa penelitian telah memanfaatkan data ulasan Google Maps untuk analisis sentimen menggunakan metode pembelajaran mesin. Misalnya, Ipmawati et al. menggunakan algoritma *Support Vector Machine* (SVM) untuk mengkategorikan sentimen ulasan wisata di Yogyakarta [2], sedangkan Aryanto dan Mardhiyyah menerapkan SVM pada ulasan pengunjung Jogja City Mall [3]. Hasil-hasil tersebut menunjukkan bahwa SVM mampu mencapai akurasi tinggi (umumnya >80–90%) dalam klasifikasi sentimen berkas teks ulasan pengguna.

Dalam sebagian besar penelitian tersebut, setiap ulasan diasumsikan memiliki satu label sentimen (positif/negatif/netral). Namun, ulasan tentang fasilitas kompleks seperti lapangan futsal mungkin mencakup beberapa aspek sekaligus (misalnya fasilitas lapangan dan pelayanan). Konsep klasifikasi *multi-label* memungkinkan satu teks diasosiasikan dengan beberapa kategori label. Pendekatan klasifikasi *multi-label* untuk analisis sentimen mikroblog dilakukan oleh Liu dan Chen [4]. Untuk klasifikasi *multi-label*, biasanya menggunakan pendekatan *One-vs-Rest* (OvR) SVM. Pendekatan ini membuat satu *classifier* untuk setiap label. Sebaliknya, SVM standar biasanya digunakan untuk klasifikasi *single-label*.

Berdasarkan latar belakang di atas, penelitian ini bertujuan untuk melakukan klasifikasi sentimen ulasan pengguna mengenai lapangan futsal di Kota Semarang menggunakan metode SVM. Data penelitian diperoleh melalui *scraping* ulasan *Google Maps* dengan ekstensi *Chrome Instant Data Scraper* dan terdiri dari 1.189 ulasan. Setiap ulasan dikategorikan ke dalam salah satu atau beberapa label aspek: Fasilitas, Pelayanan, Harga, dan Lainnya. Metode yang dikaji adalah dua pendekatan klasifikasi: (1) *single-label classification* dengan SVM dasar (setiap ulasan diberi satu label), dan (2) *multi-label classification* menggunakan SVM *One-vs-Rest* (setiap ulasan dapat memiliki beberapa label). Kinerja kedua pendekatan ini kemudian dibandingkan untuk mengetahui metode mana yang lebih efektif.

2. METODE

Tahapan yang dilakukan dalam penelitian ini digambarkan pada Gambar 1. Diagram alur penelitian menunjukkan ada 7 tahap, yakni pengumpulan data, cleaning dan pre-processing, pelabelan, pengecekan data, split data, pemodelan, dan evaluasi. Penjabaran setiap tahapan adalah sebagai berikut:



Gambar 1. Diagram Alur Penelitian

2.1 Pengumpulan Data

Data diperoleh dengan metode web scrapping pada situs *Google Maps*. *Web scrapping* adalah teknik ekstraksi data otomatis dari sebuah *website*. Data ulasan bisa digunakan untuk analisis sentimen atau tujuan lain [5].

Penelitian ini menggunakan data ulasan pengguna *Google Maps* terhadap sejumlah lapangan futsal di Kota Semarang. Data dikumpulkan menggunakan ekstensi *Chrome Instant Data Scraper* untuk ulasan dalam rentang 2022–2025. Pengambilan dilakukan secara terpisah per lapangan, lalu digabungkan menjadi satu dataset berformat CSV.

2.2 *Cleaning dan pre-processing*

Pre-processing adalah tahap awal pembelajaran mesin di mana data diubah atau dikodekan sehingga mesin dapat menemukan dan mengurai data dengan cepat [6]. Tujuan utama *pre-processing* adalah untuk menemukan, mengidentifikasi, dan mengeliminasi elemen yang tidak diperlukan demi meningkatkan kualitas data untuk performa yang lebih bagus [7]. Terdapat beberapa langkah pada tahap *cleaning* dan *pre-processing*, antara lain:

1. **Normalisasi Teks**
   Normalisasi teks dilakukan untuk mengubah teks menjadi lebih sederhana, seperti mengubah teks menjadi huruf kecil atau *lowercase*, menghapus URL, mengoreksi karakter berulang, dan menghapus karakter yang tidak relevan.
2. **Modifikasi Data**
   Modifikasi data dimaksudkan untuk mengganti kata yang tidak sesuai dan salah secara penulisan dengan kata yang sesuai dengan kaidah KBBI [8]. Modifikasi ini bertujuan untuk memperbaiki data agar meningkatkan kualitas analisis data.
3. **Tokenisasi**
   Tokenisasi adalah prosedur yang menghasilkan token dari sumber data yang didalamnya terdapat teks yang dapat dibaca oleh manusia dalam bentuk kode komputer. Mengingat bahwa tujuan dari kompilasi program adalah *source code*, metode yang ideal adalah memulai dengan *tokenizer*, yang mengubah teks menjadi urutan token, kemudian mengirimkan masing-masing token ke langkah selanjutnya. Selain itu, proses parsing melakukan validasi dan pemrosesan kode tambahan, mengubah urutan token yang telah ditokenisasi menjadi struktur data hierarkis. [9].

4. ***Stop Word Filtering***
   Penghapusan *stop words* atau konjungsi dilakukan untuk membersihkan teks dari kata-kata yang kurang berguna bagi analisis teks. Menerapkan penghapusan atau pembersihan *stop words* pada *pre-processing* dapat meningkatkan performa dan akurasi dari analisis sentimen [10].
5. ***Stemming***
   Stemming adalah proses pengurangan kata menjadi bentuk dasar atau awal kata. Misalnya kata "terjangkau" akan dikurangi menjadi "jangkau". Proses *stemming* dapat meningkatkan efisiensi algoritma karena menghapus imbuhan dari kata-kata yang memiliki makna sama tetapi memiliki variasi morfologis [11]. Pada penelitian ini, untuk memudahkan proses stemming dari teks Bahasa Indonesia, digunakan Pustaka Sastrawi. Sastrawi merupakan suatu sumber daya yang digunakan untuk memproses kata-kata menjadi bentuk aslinya [12].

2.3 Pelabelan

Setelah melalui tahap *cleaning* dan *pre-processing*, setiap ulasan akan diberi label melalui proses pelabelan. Pelabelan merupakan proses pemberian label pada karakter yang digunakan untuk mengidentifikasi variabel [13]. Pelabelan data dalam penelitian ini dibagi menjadi empat kategori, yaitu fasilitas (sarana dan prasarana), pelayanan (kualitas layanan staf atau pengelola), harga (biaya sewa dan fasilitas tambahan), serta lainnya (ulasan yang tidak termasuk dalam tiga kategori utama).

Proses pelabelan dilakukan secara semi-otomatis, yakni dengan mengkombinasikan pelabelan manual dan teknik pemrograman untuk mendeteksi kata kunci untuk meningkatkan keakuratan data [14]. Penelitian ini membandingkan dua pendekatan pelabelan, antara lain:

1. *Single-label classification*, yang menetapkan satu label untuk setiap ulasan [15].
2. *Multi-label classification*, yang memungkinkan satu ulasan memiliki lebih dari satu label sesuai dengan topik yang dibahas [15].

2.4 Pembagian Data

Setelah pelabelan data, data dibagi menjadi dua kelompok, yakni 80% untuk data latih (*train data*) dan 20% untuk data tes (*test data*). Tujuan dari pembagian ini adalah untuk memastikan bahwa model tidak hanya belajar terlalu banyak dari data pelatihan, tetapi juga mampu menggeneralisasi pola yang ditemukan untuk diterapkan pada data baru [14]. Pembagian ini diterapkan secara terpisah untuk masing-masing pendekatan pelabelan.

2.5 Ekstraksi TF-IDF

Sebelum digunakan untuk melatih model SVM, data diekstraksi terlebih dahulu menggunakan TF-IDF untuk mengubah data ulasan menjadi vektor numerik. Teknik ini membantu model menyoroti analisis kata -kata penting yang relevan dengan klasifikasi suasana hati, dan meningkatkan kinerja dalam memprediksi

model dengan tampilan teks yang lebih baik [14]. Proses ini diterapkan untuk masing-masing pendekatan pelabelan.

2.6 Implementasi Model SVM

Setelah melalui tahapan ekstraksi TF-IDF, tahap selanjutnya adalah melatih model SVM. Algoritma SVM menerapkan kernel linear untuk klasifikasi data. Proses pelatihan memungkinkan model dapat memahami pola dalam data dan memaksimalkan parameternya guna meningkatkan akurasi prediksi [14]. Penelitian ini menggunakan dua pemodelan sesuai pendekatan pelabelan yang sudah dilakukan, yakni:

1. *Single-Label Model*

   Model ini dilatih menggunakan data berlabel tunggal, di mana setiap ulasan hanya memiliki satu dari empat label yang tersedia. TF-IDF diintegrasikan dalam *pipeline* menggunakan *scikit-learn*, dan *GridSearchCV* digunakan untuk mengoptimalkan hyperparameter secara menyeluruh guna meningkatkan akurasi dan mengurangi risiko *overfitting* melalui validasi silang [16].
2. *Multi-Label Model*

   Data pelatihan model *multi-label* berasal dari pelabelan di mana satu ulasan dapat memiliki lebih dari satu label, seperti “Fasilitas” dan “Harga”. Untuk menangani hal ini digunakan pendekatan *One-vs-Rest* (OVR), yang membagi data menjadi beberapa subset biner sehingga setiap label diprediksi secara independen. Pendekatan ini efektif untuk *multi-label* karena dapat menangani banyak label sekaligus dan menangkap korelasi antar label secara implisit [17].

2.7 Evaluasi Hasil Model SVM

Evaluasi model dilakukan menggunakan data uji dengan metrik akurasi, presisi, *recall*, dan *F1-score* untuk menilai performa secara menyeluruh. Hasil dari model *single-label* dan *multi-label* dibandingkan guna menentukan pendekatan yang paling akurat untuk klasifikasi ulasan.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengumpulan Data

Penelitian ini menggunakan 1.189 data ulasan pengguna yang dikumpulkan dari 30 lapangan futsal di Kota Semarang melalui web scraping dengan ekstensi Chrome. Data mencakup ulasan selama tiga tahun terakhir dan disimpan dalam satu file gabungan. Setiap ulasan memuat informasi seperti *username*, *idreview*, *comment*, dan *time* sebagai elemen penting untuk analisis.

3.2 Hasil *Pre-Processing*

Proses *pre-proessing* data merupakan proses yang dilakukan untuk membersihkan data ulasan yang akan dianalisis dalam klasifikasi menggunakan model SVM. *Pre-processing* perlu dilakukan karena kebanyakan ulasan

menggunakan bahasa non-formal yang sulit dipahami oleh model. Tabel 1 menunjukkan sebelum dan sesudah teks melalui proses *pre-processing*.

Tabel 1. Perbandingan Teks Sebelum dan Sesudah *Pre-processing*

Sebelum *pre-processing*	**Setelah *pre-processing***
fasilitas bagus, parkir, toilet, lokasi dekat dengan jalan raya walau masuk sedikit. Bagus untuk sarana olahraga terutama pemuda remaja. Sukses prima futsal..	fasilitas bagus parkir toilet lokasi jalan raya masuk bagus sarana olahraga utama pemuda remaja sukses prima futsal
Fasilitas lapangan futsal & olahraga lain Kampus Unnes. Letaknya di FIK, Kampus Timur, yang mau ke Gaza. Tempat parkir cukup luas. Fasilitas lapangan indoor lengkap dan terawat; ada toilet, tribun yg cukup besar, dan lantai lapangan cocok untuk olahraga indoor semacam futsal, basket, dsb.	fasilitas lapang futsal olahraga kampus unnes letak fik kampus timur gaza tempat parkir luas fasilitas lapang indoor lengkap awat toilet tribun lantai lapang cocok olahraga indoor futsal basket dsb
Harga sewa lapangan murah, tapi kondisinya sangat memprihatinkan. Abis main futsal, siap-siap pulangnya sakit semua badan.	harga sewa lapang murah kondisi prihatin abis main futsal pulang sakit badan
Mantappp.. Ada mushola di tingkat 2. Ada warkopnya juga	mantap mushola tingkat 2 warkop
Rumput harus diganti karena percuma ada rumput tapi rosak	rumput ganti rumput rusak

Hasil pre-processing pada Tabel 1 dijabarkan dengan tahap-tahapnya berikut ini:

1. Normalisasi Teks

   Tahap ini melibatkan beberapa langkah normalisasi teks, seperti menghapus tanda baca (*punctuation removal*) dan mengubah seluruh huruf menjadi kecil (*case folding*). Misalnya, teks "Sukses prima futsal.." dinormalisasi menjadi "sukses prima futsal". Selain itu, karakter berulang juga dihapus, sehingga teks seperti "Mantappp.." menjadi "mantap".
2. Modifikasi Data

   Tahap ini fokus untuk mengoreksi kesalahan penulisan kata. Sebagai contoh, teks "rosak" akan menjadi "rusak" setelah dikoreksi.
3. Tokenisasi

   Teks dipisah-pisah menjadi kata-kata dengan tujuan memudahkan analisis. Sebagai contoh, teks "Tempat parkir cukup luas" menjadi "tempat parkir luas" setelah tokenisasi.
4. *Stop Word Filtering*

   Menghapus kata-kata yang kurang relevan. Sebagai contoh, kata "dan" dalam teks "dan lantai lapangan cocok" menjadi "lantai lapang cocok".
5. Stemming

   Kata-kata ditransformasikan ke dalam bentuk kata dasar. Sebagai contoh, teks "kondisinya" menjadi "kondisi" untuk mengurangi keragaman bentuk kata.

3.3 Hasil Pelabelan

Pelabelan merupakan proses pemberian kategori pada setiap ulasan guna mengidentifikasi variabel yang dianalisis. Proses ini dilakukan melalui kombinasi antara pelabelan manual dan pelabelan otomatis menggunakan pemrograman berbasis pencocokan kata kunci. Tabel 2 menunjukkan kata kunci yang digunakan untuk setiap label.

Tabel 2. Daftar Kata Kunci pada Setiap Label

Fasilitas	**Pelayanan**	**Harga**
'toilet', 'ventilasi', 'lapang', 'kondisi', 'parkir', 'lampu', 'kamar mandi', 'galon', 'kipas', 'air', 'kualitas', 'bola', 'jaring', 'lantai', 'bersih', 'bagus', 'strategis', 'nyaman', 'enak', 'wangi', 'tempat', 'fasilitas', 'rumput', 'mushola', 'musholla', 'lampu', 'cahaya'	'staf', 'responsif', 'layanan', 'admin', 'pengurus', 'abang', 'jaga', 'pesan', 'dibook', 'pelayan', 'pegawai', 'ramah', 'lucu', 'pelayanan', 'manajemen'	'toilet', 'ventilasi', 'lapang', 'kondisi', 'parkir', 'lampu', 'kamar mandi', 'galon', 'kipas', 'air', 'kualitas', 'bola', 'jaring', 'lantai', 'bersih', 'bagus', 'strategis', 'nyaman', 'enak', 'wangi', 'tempat', 'fasilitas', 'rumput', 'mushola', 'musholla', 'lampu', 'cahaya'

Label lainnya tidak memiliki daftar kata kunci karena berfungsi sebagai kategori residu yang menampung ulasan-ulasan tanpa kemunculan kata kunci dari label manapun. Untuk meningkatkan akurasi, hasil pelabelan otomatis diverifikasi secara manual oleh peneliti. Proses pelabelan dilakukan menggunakan dua pendekatan, yaitu *single-label* dan *multi-label classification*. Tabel 3 menunjukkan perbandingan hasil pelabelan single-label dan multi-label.

Tabel 3. Perbandingan *Single-Label* dan *Multi-Label*

Ulasan sebelum *pre-processing*	***Single-Label***	***Multi-Label***			
		Fasilitas	Pelayanan	Harga	Lainnya
Murah ada parkir nya, ada kamar mandi yg bagus	Fasilitas	1	0	1	0
Tempat nyaman, admin baik & ramah, remonded	Fasilitas	1	1	0	0
Main futsal mantap	Lainnya	0	0	0	1

1. *Single-Label*

   Pendekatan ini menggunakan *single-label classification* berbasis pencocokan kata kunci, di mana label ditetapkan berdasarkan jumlah kata kunci terbanyak dalam ulasan. Misalnya, “Murah ada parkirnya, ada kamar mandi yang bagus” memiliki lebih banyak kata kunci fasilitas, sehingga diberi label fasilitas. Jika jumlahnya sama, label yang muncul lebih awal dipilih, seperti pada “Tempat nyaman, admin baik & ramah, recommended”

yang diberi label fasilitas. Ulasan tanpa kata kunci, seperti "Main futsal mantap", dikategorikan sebagai Lainnya.

2. *Multi-Label*

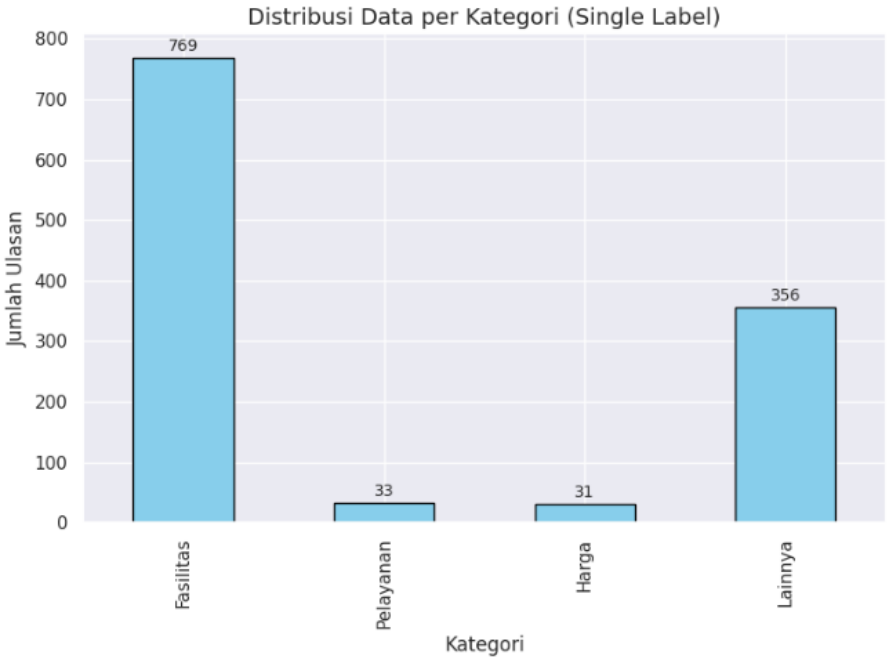
   Pendekatan ini memungkinkan satu ulasan memiliki lebih dari satu label, dengan tiap label direpresentasikan dalam format biner: 1 jika mengandung kata kunci relevan, dan 0 jika tidak. Contohnya, ulasan "Murah ada parkirnya, ada kamar mandi yg bagus" diberi nilai 1 untuk fasilitas dan harga, serta 0 untuk pelayanan dan lainnya. Ulasan "Tempat nyaman, admin baik & ramah, recommended" diberi nilai 1 untuk fasilitas dan pelayanan. Sedangkan ulasan "Main futsal mantap" tidak mengandung kata kunci dari ketiga label utama, sehingga hanya diberi nilai 1 untuk label lainnya.

3.4 Hasil Klasifikasi Ulasan Menggunakan SVM

Model Support Vector Machine (SVM) dilatih dengan data ulasan yang telah melalui *pre-processing* seperti pembersihan teks dan *stemming*. Fitur diekstraksi menggunakan TF-IDF, kemudian model dibangun secara terpisah untuk skenario *single-label* dan *multi-label* sesuai pendekatannya.

1. Pemodelan *Single-Label*

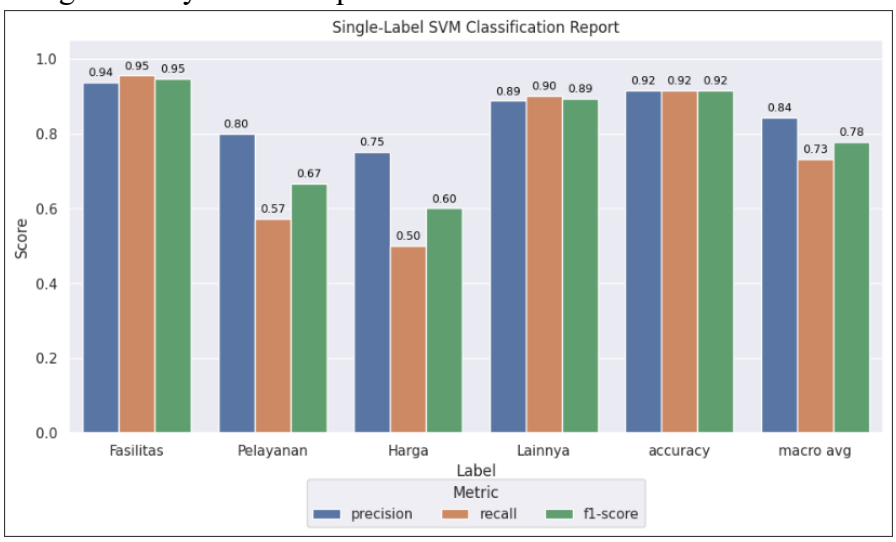
   Model single-label dilatih dengan data berlabel tunggal per ulasan, dan optimasi hiperparameternya dilakukan menggunakan *GridSearchCV* untuk memperoleh kombinasi parameter terbaik guna meningkatkan performa klasifikasi.



Gambar 2. Hasil Sebaran Data *Single-Label*

Gambar 2 menunjukkan distribusi ulasan pengguna untuk klasifikasi *single-label* yang tidak seimbang. Kategori fasilitas mendominasi dengan 769 ulasan, jauh lebih banyak dibandingkan kategori pelayanan dan harga

yang masing-masing hanya memiliki 33 dan 31 ulasan. Sementara itu, kategori lainnya mencakup 356 ulasan.



Gambar 3. Hasil Metrik Performa Analisis *Single-Label*

Hasil evaluasi model SVM ditampilkan pada gambar 3, yang meliputi beberapa metrik evaluasi. Berdasarkan hasil tersebut, beberapa penemuan utama meliputi:

a. Presisi

Nilai presisi tertinggi diperoleh pada label fasilitas sebesar 0.94, menunjukkan bahwa sebagian besar prediksi label ini relevan. Label pelayanan memiliki presisi 0.80, cukup baik meski datanya terbatas. Presisi terendah terdapat pada label harga sebesar 0.75, menandakan masih terdapat 25% kesalahan klasifikasi. Sementara itu, label lainnya mencatat presisi 0.89, mencerminkan kinerja model yang cukup andal dalam mengidentifikasi ulasan di luar tiga kategori utama.

b. *Recall*

Label fasilitas mencatat *recall* tertinggi sebesar 0.95, menunjukkan kemampuan model yang sangat baik dalam mengenali ulasan relevan. Sebaliknya, *recall* label pelayanan hanya 0.57, menandakan banyak ulasan yang terlewat. *Recall* terendah terdapat pada label harga sebesar 0.50, menunjukkan hanya separuh ulasan yang berhasil diidentifikasi. Sementara itu, label lainnya memiliki *recall* 0.90, mencerminkan performa model yang kuat dalam mengenali ulasan di luar tiga kategori utama.

c. *F1-Score*

Label fasilitas meraih *F1-score* tertinggi sebesar 0.95, menunjukkan keseimbangan optimal antara presisi dan *recall*. Label pelayanan memperoleh *F1-score* 0.67, mencerminkan performa moderat dalam mengenali ulasan relevan. Label harga mencatat *F1-score* terendah sebesar 0.60, menandakan tantangan dalam klasifikasi ulasan pada kategori ini.

Sementara itu, label lainnya mencapai *F1-score* 0.89, menunjukkan kinerja yang cukup baik dalam mendeteksi ulasan di luar tiga kategori utama.

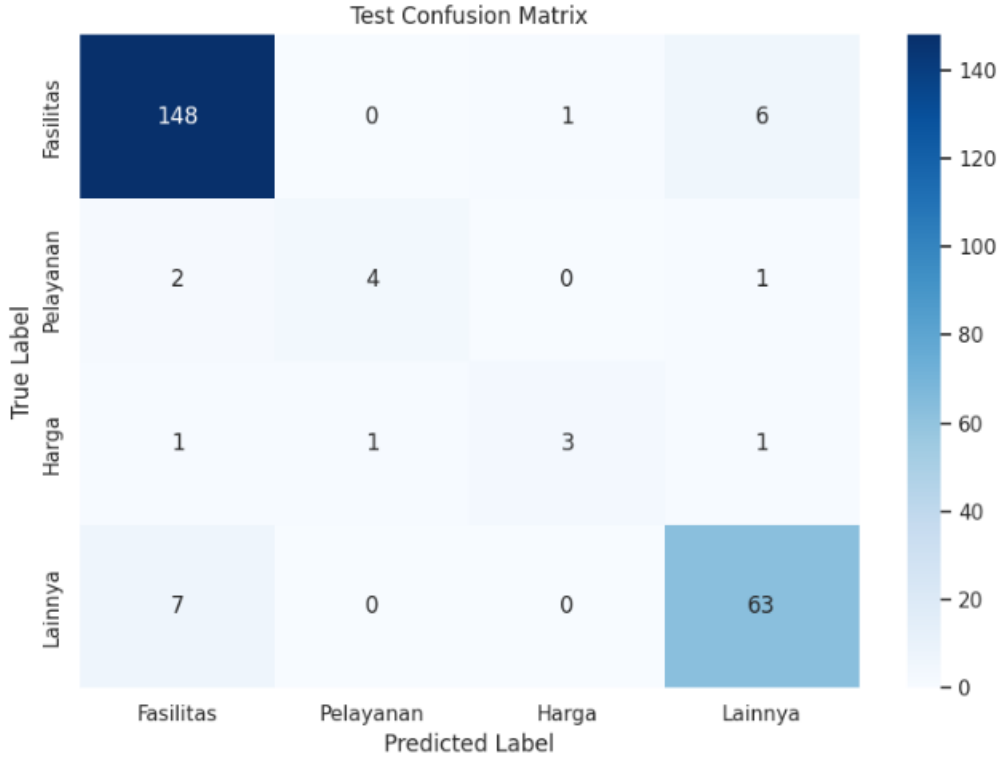
d. *Support*

Distribusi ulasan yang tidak seimbang, dengan dominasi label fasilitas (155 ulasan) dibandingkan pelayanan (7 ulasan), harga (6 ulasan), dan lainnya (70 ulasan), menyebabkan model lebih terlatih pada kelas mayoritas. Ketimpangan ini berdampak pada rendahnya nilai *recall* dan *F1-score* untuk kategori minoritas, mengindikasikan bahwa model mengalami kesulitan dalam mengenali pola pada kelas dengan jumlah data yang terbatas.

e. Akurasi

Model mencapai akurasi keseluruhan 0.92, menunjukkan 92% prediksi sesuai dengan label sebenarnya. Namun, akurasi ini harus diinterpretasikan hati-hati karena ketidakseimbangan data antar kategori. Pada kondisi *class imbalance*, akurasi bisa kurang representatif terutama dalam menilai performa pada kelas minoritas dengan data terbatas.

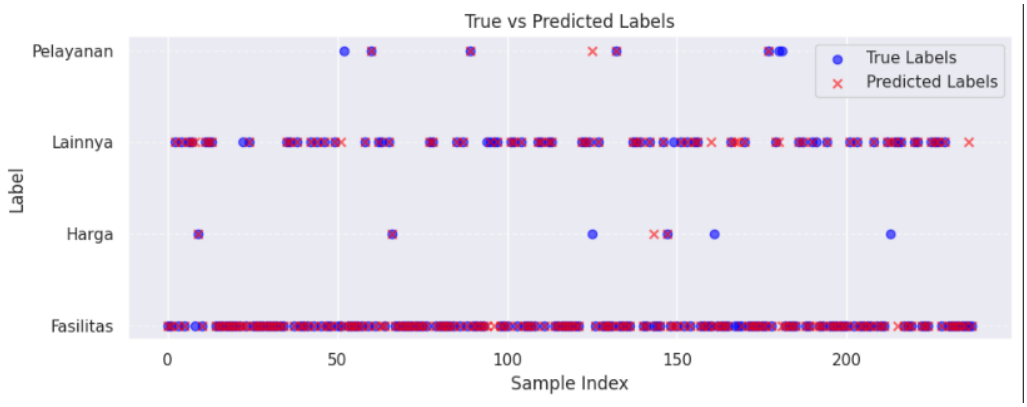
f. *Macro Average*: Nilai presisi rata-rata sebesar 0.84, *recall* rata-rata 0.73, dan *F1-score* rata-rata 0.78 menunjukkan bahwa model secara keseluruhan memiliki performa yang cukup baik, meskipun kinerja antar kategori bervariasi akibat ketidakseimbangan jumlah data.



Gambar 4. Hasil *Confussion Matrix* Analisis Single-Label

Gambar 4 memperlihatkan *confusion matrix* hasil prediksi model SVM untuk empat label: fasilitas, pelayanan, harga, dan lainnya. Model

menunjukkan performa baik terutama pada kategori dengan data dominan. Label fasilitas terklasifikasi benar sebanyak 148 ulasan, dengan 1 ulasan salah ke harga dan 6 ke lainnya. Label pelayanan benar 4 ulasan, dengan 2 salah ke fasilitas dan 1 ke Lainnya. Label harga memiliki akurasi terendah, hanya 3 ulasan benar, sementara masing-masing 1 ulasan salah diklasifikasikan ke fasilitas, pelayanan, dan lainnya. Labrl lainnya benar 63 ulasan, dengan 7 salah ke Fasilitas. Ketidakseimbangan data berkontribusi pada kesalahan klasifikasi terutama pada kelas minoritas.

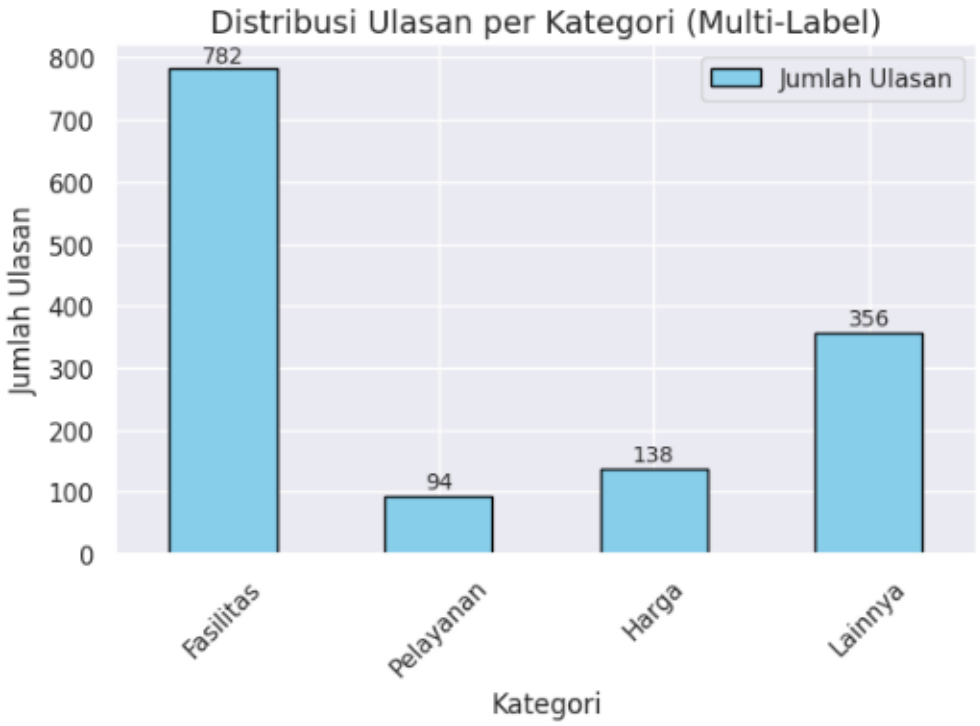


Gambar 5. Hasil Grafik Distribusi Analisis *Single-Label*

Gambar 5 memperlihatkan perbandingan label aktual (lingkaran biru) dan prediksi (X merah) pada analisis *single-label*. Prediksi untuk label fasilitas hampir sepenuhnya tepat, sedangkan label lainnya sebagian besar sesuai dengan label asli. Namun, prediksi untuk pelayanan dan harga banyak yang meleset, kemungkinan karena jumlah data yang terbatas. Secara keseluruhan, model SVM sangat baik pada label fasilitas, cukup pada label lainnya, dan kurang akurat pada label pelayanan dan harga.

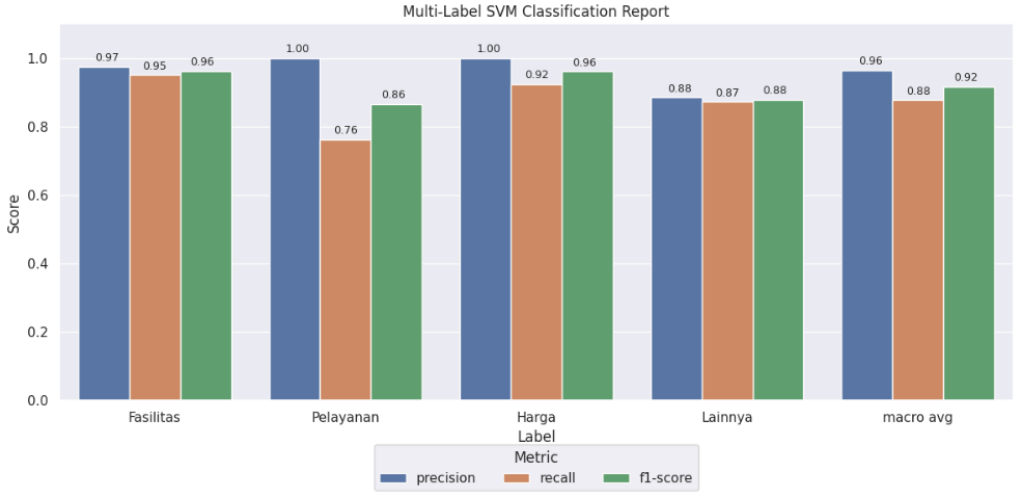
2. Pemodelan *Multi-Label*

Model klasifikasi *multi-label* pada penelitian ini menggunakan pendekatan *One-vs-Rest* (OvR), di mana setiap label seperti fasilitas, pelayanan, dan harga dipelajari secara terpisah melalui pelatihan biner. Setiap ulasan dievaluasi terhadap masing-masing label secara independen tanpa mempertimbangkan label lain.



Gambar 6. Hasil Sebaran Data *Multi-Label*

Gambar 6 menunjukkan distribusi ulasan pengguna dalam klasifikasi *multi-label*, dengan label fasilitas sebagai yang mendominasi, yaitu sebanyak 782 ulasan. Adapun label pelayanan dan harga masing-masing hanya mencakup 94 dan 138 ulasan. Sedangkan label lainnya mencakup 356 ulasan. Ketimpangan ini menunjukkan ketidakseimbangan kelas yang dapat memengaruhi performa model klasifikasi.



Gambar 7. Hasil Metrik Performa Analisis *Multi-Label*

Gambar 7 menunjukkan hasil evaluasi model SVM, yang meliputi beberapa metrik evaluasi diuraikan sebagai berikut:

a. Presisi

Nilai presisi model SVM menunjukkan performa tinggi. Label fasilitas memiliki presisi 0,97, menandakan prediksi yang hampir seluruhnya tepat. Label pelayanan dan harga masing-masing mencapai presisi 1,00, artinya tidak ada kesalahan dalam pelabelan. Sementara itu, label lainnya memperoleh presisi 0,88, menunjukkan model cukup akurat dalam mengklasifikasikan ulasan di luar tiga kategori utama.

b. *Recall*

Recall tertinggi dicapai pada label fasilitas sebesar 0,95, yang menunjukkan bahwa sebagian besar ulasan terkait berhasil dikenali model. Label harga dan lainnya juga menunjukkan kinerja baik dengan recall masing-masing 0,92 dan 0,87. Sementara itu, label pelayanan memiliki recall terendah sebesar 0,76, menandakan adanya ulasan relevan yang tidak terdeteksi meskipun presisinya sempurna, sehingga model cenderung melewatkan beberapa data yang semestinya terklasifikasi.

c. *F1-Score*

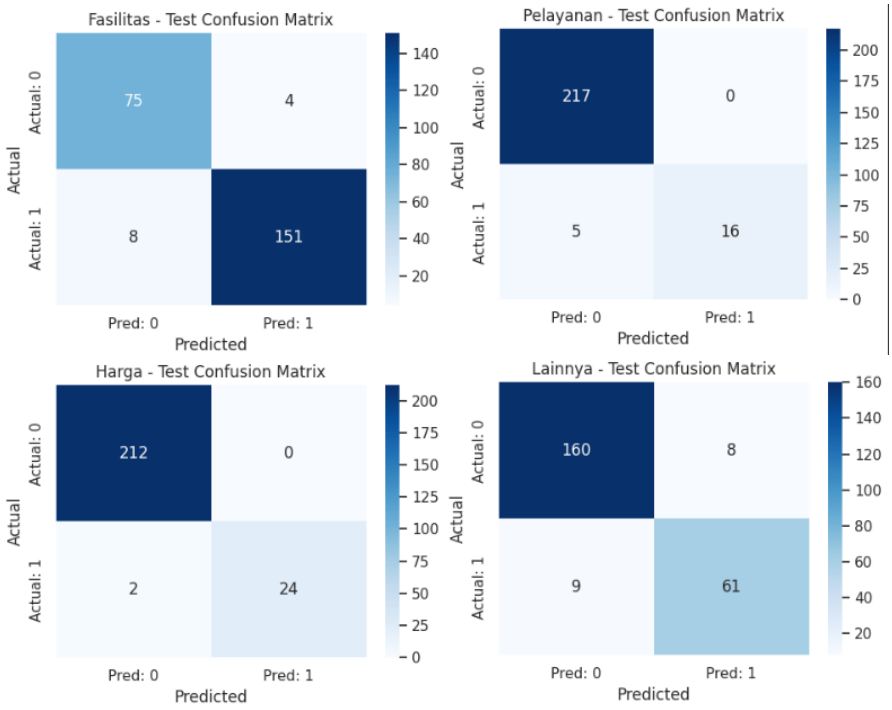
F1-Score tertinggi diperoleh label fasilitas (0,96), disusul harga (0,92) dan lainnya (0,88), menunjukkan performa stabil. Label pelayanan memiliki *F1-Score* terendah (0,86) akibat presisi tinggi tetapi recall rendah, menandakan banyak ulasan relevan yang tidak terdeteksi.

d. *Support*

Dalam pendekatan *multi-label*, satu ulasan dapat memiliki lebih dari satu label, sehingga total *support* mencapai 276 meskipun jumlah data uji hanya 238. Distribusi label juga tidak seimbang, dengan label fasilitas sebanyak 159 ulasan, pelayanan 21, harga 26, dan lainnya 70. Ketimpangan ini memengaruhi performa model, yang cenderung lebih baik pada kategori mayoritas dan menurun pada kategori minoritas, terlihat dari rendahnya nilai *recall* dan *f1-score*.

e. *Macro Average*

Nilai *macro average* menunjukkan performa keseluruhan model: presisi 0,96; *recall* 0,88; dan *F1-score* 0,92. Secara umum, model bekerja cukup baik, namun masih terdapat variasi kinerja antar kategori akibat ketidakseimbangan jumlah data.



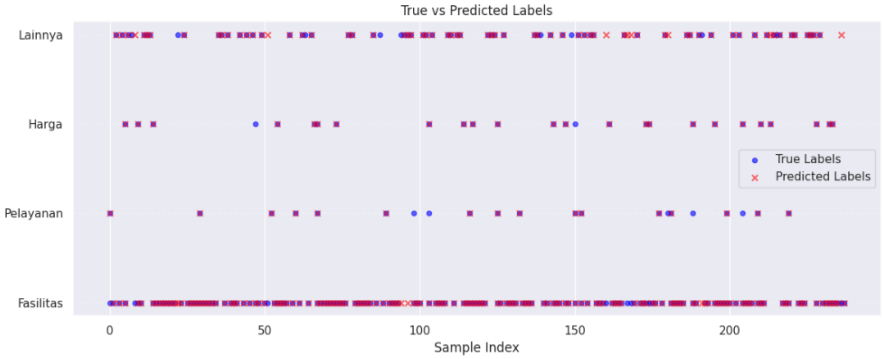
Gambar 8. Hasil *Confussion Matrix* Analisis *Multi-Label*

Gambar 8 menampilkan *confusion matrix* hasil pemodelan SVM dengan pendekatan *One-vs-Rest* (OvR). Dalam pendekatan ini, model dibangun secara terpisah untuk tiap kategori guna membedakan antara data yang termasuk dan tidak termasuk dalam label tersebut.

Confusion matrix dalam pemodelan memiliki dua sumbu utama, yaitu sumbu vertikal yang merepresentasikan label aktual dan sumbu horizontal yang menunjukkan hasil prediksi model. Nilai 1 menandakan ulasan termasuk dalam kategori, sedangkan 0 berarti tidak termasuk. Kombinasi nilai ini membentuk empat kemungkinan: "*True Positive*" ketika ulasan benar terklasifikasi, "*False Negative*" saat ulasan relevan tidak dikenali, "*False Positive*" ketika model salah mengklasifikasikan ulasan yang tidak relevan, dan "*True Negative*" saat model dengan benar mengenali ulasan yang tidak termasuk dalam kategori.

Pada label fasilitas, model mencatat 151 *true positive* dan hanya 4 *false positive*, menghasilkan presisi tinggi sebesar 0.97. Meskipun terdapat 8 *false negative*, *recall* tetap tinggi pada angka 0.95. Untuk label pelayanan, presisi mencapai 1.00 karena tidak ada *false positive*, namun *recall* menurun menjadi 0.76 akibat 5 ulasan yang tidak terdeteksi dari total 21. Label harga juga menunjukkan presisi sempurna, dengan *recall* 0.92 dari 24 ulasan yang terdeteksi dari total 26. Sementara itu, pada label lainnya, model menghasilkan presisi 0.88 dan *recall* 0.87, dengan 61 *true positive*, 8 *false positive*, dan 9 *false negative*. Secara umum, model menunjukkan performa

baik, terutama dalam mengenali ulasan positif secara tepat dan menjaga kestabilan terhadap data negatif.



Gambar 9. Hasil Grafik Distribusi Analisis Multi-Label

Gambar 9 memperlihatkan perbandingan *true-label* dan *predicted-label*. Label fasilitas dan lainnya menunjukkan prediksi yang hampir sepenuhnya akurat, sedangkan label pelayanan dan harga memiliki beberapa kesalahan prediksi, kemungkinan karena jumlah data yang lebih sedikit, sehingga kinerja model kurang optimal pada label tersebut.

3.5 Diskusi

Berdasarkan hasil analisis, model SVM multi-label menunjukkan performa yang lebih baik dibandingkan model *single-label*, khususnya pada metrik presisi dan *F1-score* makro rata-rata (*multi-label*: presisi 0.96, *F1-score* 0.92; *single-label*: presisi 0.84, *F1-score* 0.78). Perbedaan ini muncul karena pendekatan *multi-label* mampu menangani ulasan yang memiliki lebih dari satu aspek sekaligus, sedangkan *single-label* hanya mengklasifikasikan ulasan ke satu kategori dominan sehingga mengabaikan informasi lain.

Dari sisi performa per-kategori, model *single-label* unggul di label fasilitas yang memiliki data dominan (769 data), dengan presisi 0.94 dan *F1-score* 0.95. Namun, performanya turun tajam di kategori minor seperti label pelayanan dan harga, yang masing-masing hanya memiliki 33 dan 31 data (*F1-score* 0.57 dan 0.60). Ini mengindikasikan bahwa model *single-label* sangat rentan terhadap ketidakseimbangan data. Sebaliknya, model *multi-label* mampu mempertahankan performa tinggi di seluruh label, termasuk kategori minor (pelayanan dan harga dengan *F1-score* 0.86 dan 0.96), berkat strategi *One-vs-Rest* yang mengurangi dampak ketidakseimbangan dengan memodelkan setiap label secara independen.

Ketidakseimbangan data sangat mempengaruhi performa kedua model. Model *single-label* bias terhadap kategori mayoritas, menyebabkan rendahnya *recall* dan presisi di kategori minor. Sementara model *multi-label* lebih adaptif karena memperlakukan setiap label secara terpisah, sehingga lebih tahan terhadap distribusi data yang tidak merata.

Untuk meningkatkan performa, khususnya model *single-label*, dapat diterapkan beberapa strategi seperti oversampling label minor (misal SMOTE), pemberian bobot kelas (*class weight*) selama pelatihan, penggunaan model *hybrid* yang menggabungkan keunggulan *single-label* dan *multi-label*, serta augmentasi data berbasis teks untuk memperkaya data minor tanpa perlu pengumpulan data manual.

Secara metodologis dan praktis, perbedaan kinerja model *single-label* dan *multi-label* sangat signifikan. Secara metodologis, hasil ini menunjukkan bahwa strategi klasifikasi yang dipilih sangat memengaruhi kemampuan model untuk menangkap kompleksitas ulasan multi-aspek. Dalam konteks ulasan lapangan futsal, di mana satu ulasan sering membahas lebih dari satu aspek sekaligus, pendekatan *multi-label* mampu mempertahankan informasi yang hilang pada pendekatan *single-label*, sehingga memberikan gambaran yang lebih komprehensif. Pendekatan *multi-label* juga menjelaskan kemampuan multi-label dalam menjaga penilaian yang lebih akurat.

Dari perspektif praktis, penggunaan model yang tepat akan berdampak pada tingkat pemahaman yang diperoleh pengelola lapangan futsal. Meskipun model *single-label* lebih sederhana, ia mungkin menghasilkan analisis yang bias ke kategori mayoritas dan mengabaikan masalah penting pada kategori minor. Sebaliknya, model *multi-label* dapat menghasilkan ulasan yang mencakup berbagai elemen, seperti fasilitas dan pelayanan, sehingga manajer dapat membuat intervensi yang lebih tepat sasaran. Temuan ini menunjukkan bahwa pemilihan pendekatan multi-label berdampak langsung pada efisiensi pengambilan keputusan berbasis data, bukan hanya pertimbangan teknis.

Secara keseluruhan, pendekatan SVM *multi-label* dengan *One-vs-Rest* lebih efektif dan direkomendasikan untuk klasifikasi ulasan pengguna yang kompleks dan beragam, terutama pada data yang tidak seimbang seperti ulasan lapangan futsal. Temuan ini tidak hanya memberikan solusi praktis untuk pengelola lapangan, tetapi juga berkontribusi pada pengembangan metode klasifikasi ulasan berbasis machine learning dalam domain analisis opini, khususnya dalam menangani data *multi-label* dengan distribusi yang tidak merata.

3.6 Kebaruan Penelitian

Selain membahas perbedaan kinerja model, penting untuk mengaitkan penelitian ini dengan studi sebelumnya untuk menegaskan kontribusi. Ipmawati et al. (2024) menggunakan algoritma SVM untuk menganalisis sentimen ulasan *Google Maps* pada objek wisata Yogyakarta [2]. Pendekatan *single-label* digunakan untuk penelitian ini, yang memiliki tiga kategori sentimen (positif, negatif, dan netral) dan berfokus pada evaluasi akurasi model tanpa mempertimbangkan kemungkinan bahwa satu ulasan mengandung lebih dari satu aspek. Serupa dengan ini, Aryanto dan Mardhiyyah (2024) menganalisis sentimen ulasan *Google Maps pada* pusat perbelanjaan (Jogja City Mall), menggunakan SVM *single-label*, dengan klasifikasi polaritas dan tanpa menguji model *multi-label* [3]. Kedua penelitian ini tidak membahas dampak ketidakseimbangan distribusi label *multi-label* maupun strategi klasifikasi yang mampu menangani lebih dari satu label pada satu entri ulasan.

Tidak seperti penelitian sebelumnya, penelitian ini memeriksa ulasan *Google Maps* lapangan futsal di Indonesia dengan membandingkan dua metode klasifikasi: *single-label* dan *multi-label*. Untuk klasifikasi *multi-label*, strategi *One-vs-Rest* (OvR) digunakan. Selain itu, untuk meningkatkan kualitas data, penelitian ini menggabungkan teknik pelabelan semi-otomatis berbasis kata kunci yang diverifikasi secara manual dengan menggunakan fase pemrosesan bahasa alami seperti normalisasi, tokenisasi, *stop word filtering*, dan *stemming*. Selain itu, penelitian ini melihat bagaimana ketidakseimbangan distribusi label *multi-label* mempengaruhi kinerja model. Hal ini belum pernah dibahas dalam studi

sebelumnya tentang fasilitas olahraga. Metode ini menawarkan kontribusi baru dalam pengembangan metode analisis sentimen berbasis SVM untuk data multi-aspek dengan distribusi label yang tidak seimbang.

3.7 Implikasi Praktis

Temuan penelitian ini memiliki hubungan langsung dengan manajemen lapangan futsal, khususnya dalam hal mengidentifikasi dan memenuhi permintaan pengguna dengan lebih cepat dan tepat. Keunggulan model *multi-label* SVM, yang dapat menangkap lebih dari satu elemen pada setiap ulasan, memungkinkan pengelola untuk mendapatkan gambaran menyeluruh tentang bagaimana pengguna melihat fasilitas, layanan, dan harga. Keputusan yang berbasis data dapat dibuat dengan data ini, misalnya memperbaiki fasilitas yang paling sering dikritik atau meningkatkan aspek pelayanan yang dianggap kurang memuaskan.

Model *multi-label* SVM yang dikembangkan dapat diintegrasikan ke dalam platform pemesanan online yang sudah digunakan pengelola atau sistem informasi manajemen lapangan futsal. Modul analisis sentimen memudahkan integrasi ini dengan mengumpulkan data ulasan dari *Google Maps* secara otomatis, memprosesnya, dan menampilkan hasil analisis dalam dashboard interaktif. Per-aspek, dashboard ini dapat menampilkan indikator tren sentimen, peringatan dini penurunan kualitas, dan saran untuk tindakan lanjut. Dengan sistem seperti ini, pengelola tidak perlu membaca ulasan secara manual, yang mengurangi waktu respons terhadap masalah dan meningkatkan efisiensi operasi.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa model *multi-label* SVM secara konsisten memberikan performa klasifikasi yang lebih baik dan merata dibandingkan model *single-label*, terutama dalam konteks ulasan yang mengandung banyak aspek sekaligus. Perbedaan performa ini sangat dipengaruhi oleh ketidakseimbangan data yang cukup signifikan pada kategori minor seperti label pelayanan dan harga, yang menyebabkan model *single-label* kesulitan mengenali kategori tersebut secara optimal. Sebaliknya, pendekatan *multi-label* dengan strategi *One-vs-Rest* terbukti lebih adaptif terhadap distribusi data tidak seimbang dan mampu menangani kompleksitas ulasan *multi-label* dengan lebih baik.

Meskipun *multi-label* SVM menunjukkan keunggulan, implementasinya perlu memperhatikan aspek kompleksitas model dan waktu pelatihan yang lebih tinggi. Sementara itu, pendekatan *single-label* tetap relevan apabila data cukup seimbang dan ulasan cenderung fokus pada satu aspek utama.

Untuk meningkatkan performa, terutama pada model *single-label*, beberapa strategi seperti *oversampling* kategori minor (contohnya dengan SMOTE), pemberian bobot kelas selama pelatihan, augmentasi data berbasis teks, serta pengembangan model *hybrid* dapat diadopsi.

Hasil penelitian ini tidak hanya memberikan kontribusi praktis untuk pengelolaan layanan publik seperti lapangan futsal, tetapi juga memperkaya kajian klasifikasi *multi-label* dalam bidang analisis opini berbasis *machine learning*. Dengan pemilihan metode yang tepat, sistem klasifikasi ulasan dapat menjadi alat yang efektif dalam memahami dan merespon kebutuhan pengguna secara lebih komprehensif.

REFERENSI

[1] A. N. Putra and S. Soegiyanto, "Manajemen Pembinaan Prestasi Futsal Kota Semarang dalam Persiapan Menghadapi PORPROV Jawa Tengah Tahun 2023," *Jurnal Sains Keolahragaan dan Kesehatan*, vol. 8, no. 2, pp. 161–176, Mar. 2024, doi: 10.5614/jskk.2023.8.2.6.

[2] J. Ipmawati, S. Saifulloh, and K. Kusnawi, "Analisis Sentimen Tempat Wisata Berdasarkan Ulasan pada Google Maps Menggunakan Algoritma Support Vector Machine," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 247–256, Jan. 2024, doi: 10.57152/malcom.v4i1.1066.

[3] P. M. Aryanto and R. Mardhiyyah, "Analisis Sentimen Terhadap Review Google Maps Jogja City Mall Menggunakan Algoritma Support Vector Machine," 2024, doi: 10.47065/josyc.v6i1.6049.

[4] S. M. Liu and J.-H. Chen, "A multi-label classification based approach for sentiment classification," *Expert Syst Appl*, vol. 42, no. 3, pp. 1083–1093, Feb. 2015, doi: 10.1016/j.eswa.2014.08.036.

[5] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowl Based Syst*, vol. 226, p. 107134, Aug. 2021, doi: 10.1016/j.knosys.2021.107134.

[6] J. Hartmann, M. Heitmann, C. Siebert, and C. Schamp, "More than a Feeling: Accuracy and Application of Sentiment Analysis," *International Journal of Research in Marketing*, vol. 40, no. 1, pp. 75–87, Mar. 2023, doi: 10.1016/j.ijresmar.2022.05.005.

[7] A. M. S. Shaik Afzal, "Optimized Support Vector Machine Model for Visual Sentiment Analysis," in *2021 3rd International Conference on Signal Processing and Communication (ICPSC)*, IEEE, May 2021, pp. 171–175. doi: 10.1109/ICSPC51351.2021.9451669.

[8] B. Irena and Erwin Budi Setiawan, "Fake News (Hoax) Identification on Social Media Twitter using Decision Tree C4.5 Method," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 4, pp. 711–716, Aug. 2020, doi: 10.29207/resti.v4i4.2125.

[9] R. Friedman, "Tokenization in the Theory of Knowledge," *Encyclopedia*, vol. 3, no. 1, pp. 380–386, Mar. 2023, doi: 10.3390/encyclopedia3010024.

[10] N. Umar and M. Adnan Nur, "Application of Naïve Bayes Algorithm Variations On Indonesian General Analysis Dataset for Sentiment Analysis," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 4, pp. 585–590, Aug. 2022, doi: 10.29207/resti.v6i4.4179.

[11] M. W. Sardjono, M. Cahyanti, M. Mujahidin, and R. Arianty, "PENDETEKSI KESAMAAN KATA UNTUK JUDUL PENULISAN BERBAHASA INDONESIA MENGGUNAKAN ALGORITMA STEMMING NAZIEF-ADRIANI," *Sebatik*, vol. 22, no. 2, pp. 138–146, Dec. 2018, doi: 10.46984/sebatik.v22i2.320.

[12] B. Siswanto and Y. Dani, "Sentiment Analysis about Oximeter as Covid-19 Detection Tools on Twitter Using Sastrawi Library," in *2021 8th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE)*, IEEE, Sep. 2021, pp. 161–164. doi: 10.1109/ICITACEE53184.2021.9617216.

[13] Z. Rais, F. T. T. Hakiki, and R. Aprianti, "Sentiment Analysis of Peduli Lindungi Application Using the Naive Bayes Method," *SAINSMAT: Journal of Applied Sciences, Mathematics, and Its Education*, vol. 11, no. 1, pp. 23–29, Mar. 2022, doi: 10.35877/sainsmat794.

[14] N. Widaad and D. Anggraini, "Sentiment Analysis Of Chatgpt App User Reviews Using Svm And Cnn Methods," *Jurnal Teknik Informatika (JUTIF)*, vol. 5, no. 6, pp. 1687–1700, 2024, doi: 10.52436/1.jutif.2024.5.6.4010.

[15] N. A. Sajid *et al.*, "Single vs. Multi-Label: The Issues, Challenges and Insights of Contemporary Classification Schemes," Jun. 01, 2023, *MDPI*. doi: 10.3390/app13116804.

[16] K. Alemerien, S. Alsarayreh, and E. Altarawneh, "Diagnosing Cardiovascular Diseases using Optimized Machine Learning Algorithms with GridSearchCV," *Journal of Applied Data Sciences*, vol. 5, no. 4, pp. 1539–1552, Dec. 2024, doi: 10.47738/jads.v5i4.280.

[17] J. Xu, "An extended one-versus-rest support vector machine for multi-label classification," *Neurocomputing*, vol. 74, no. 17, pp. 3114–3124, Oct. 2011, doi: 10.1016/j.neucom.2011.04.024.