

Terakreditasi SINTA Peringkat 3

Surat Keputusan Direktur Jenderal Pendidikan Tinggi, Riset, dan Teknologi Nomor 225/E/KPT/2022 masa berlaku mulai Vol.7 No. 1 tahun 2022 s.d Vol. 11 No. 2 tahun 2026

Terbit online pada laman web jurnal:
<http://publishing-widyagama.ac.id/ejournal-v2/index.php/jointecs>



Vol. 8 No. 3 (2023) 105 - 114

JOINTECS

(Journal of Information Technology and Computer Science)

e-ISSN:2541-6448

p-ISSN:2541-3619

Gesture Recognition untuk Deteksi Bahasa Isyarat BISINDO: Pendekatan Mediapipe dan Random Forest

Salsabila Ayuni Kaffah¹, Yudi Ramdhani^{2*}

Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Adhirajasa Reswara Sanjaya

¹salsaakffha@gmail.com, ^{2*}yudi@ars.ac.id

Abstract

Gesture Recognition plays a crucial role in facilitating and enhancing communication accessibility for individuals with hearing and speech impairments. However, translating complex sign language into spoken or written language remains a significant challenge. In an effort to address this, this research utilizes the MediaPipe framework and the Random Forest Classifier algorithm to classify sign language gestures and expressions in BISINDO (Indonesian Sign Language). Considering the difficulty and complexity of sign language gestures, 10 expressions/words in BISINDO were selected, resulting in a dataset of 25,000 data points used in this study. The approach involves detecting sign language through pose, hand, and facial gesture or movement recognition. Evaluation results show that the Random Forest algorithm achieves a remarkably high level of precision, recall, F1-score, and accuracy (99.88%). Additionally, the developed system demonstrates good performance with prediction probabilities ranging from 0.50 to 0.70 for correct predictions, although challenges persist in distinguishing similar sign gestures, resulting in some predictions requiring more time to yield accurate results. The findings of this research contribute significantly to improving sign language recognition and promoting inclusivity for individuals with hearing and speech impairments. Moreover, it opens up new opportunities for further advancements in sign language detection technology.

Keywords: BISINDO Sign Language; Hearing Impairment; Machine Learning; MediaPipe; Random Forest.

Abstrak

Gesture Recognition memainkan peran penting dalam memfasilitasi dan meningkatkan aksesibilitas komunikasi bagi individu dengan gangguan pendengaran dan bicara. Namun, dalam menerjemahkan bahasa isyarat yang kompleks menjadi bahasa lisan atau tulisan tetap menjadi tantangan yang signifikan. Berupaya untuk mengatasi hal tersebut, penelitian ini memanfaatkan framework MediaPipe dan algoritma Random Forest Classifier untuk mengklasifikasikan gerakan isyarat berbentuk ungkapan dan kata dalam bahasa isyarat BISINDO. Dengan mempertimbangkan tingkat kesulitan dan kompleksitas gerakan isyarat, 10 label ungkapan/kata dalam BISINDO dipilih dan menghasilkan total 25.000 data yang dipakai pada sistem di penelitian ini. Pendekatan ini melibatkan deteksi bahasa isyarat melalui pengenalan pose, gerakan tangan, dan ekspresi wajah. Hasil evaluasi menunjukkan algoritma Random Forest mencapai tingkat presisi, recall, F1-score, dan akurasi yang sangat tinggi (99,88%). Selain itu, sistem yang dikembangkan juga menunjukkan kinerja baik dengan rata-rata probabilitas prediksi berkisar antara 0,50 hingga 0,70 untuk prediksi yang benar, meskipun terdapat tantangan dalam membedakan gerakan isyarat yang mirip dan menyebabkan beberapa prediksi memerlukan waktu lebih lama untuk mencapai hasil yang tepat. Dengan hasil yang diperoleh, penelitian ini memberikan kontribusi penting dalam meningkatkan pengenalan bahasa isyarat dan mendorong inklusivitas bagi masyarakat dengan gangguan pendengaran dan bicara. Hal ini juga membuka peluang baru untuk pengembangan lebih lanjut dalam teknologi deteksi bahasa isyarat.

Kata kunci: Bahasa Isyarat BISINDO; Gangguan Pendengaran; Machine Learning; MediaPipe; Random Forest.



Diterima Redaksi : 14-07-2023 | Selesai Revisi : 23-07-2023 | Diterbitkan Online : 30-09-2023

1. Pendahuluan

Menurut WHO (*World Health Organization*), saat ini lebih dari 5% dari populasi dunia (432 juta) hidup dengan gangguan pendengaran termasuk 34 juta anak-anak. Diperkirakan pada tahun 2050 lebih dari 700 juta orang – atau 1 dari setiap 10 orang – akan mengalami gangguan pendengaran [1]. Bahasa merupakan sarana komunikasi yang digunakan oleh individu baik dalam kelompok sosial maupun budayanya untuk menyampaikan informasi dan pendapat kepada orang lain [2]. Namun, orang dengan gangguan pendengaran (Tuna Rungu) dan bicara (Tuna Wicara) tidak dapat menggunakan bahasa tersebut untuk berkomunikasi, untuk itu mereka menggunakan cara lain yang disebut dengan bahasa isyarat. Bahasa Isyarat dihasilkan oleh gerak tangan, ekspresi wajah, dan tubuh bagian atas, lalu diproses oleh mata [3]. Bahasa isyarat dikembangkan untuk memfasilitasi komunikasi yang efektif bagi orang-orang yang mengalami gangguan pendengaran dan bicara. Tidak ada bentuk bahasa isyarat universal yang telah disepakati, sehingga bahasa isyarat dari negara atau daerah yang berbeda tidak dapat saling dimengerti [4]. Di Indonesia, terdapat dua penggunaan bahasa isyarat yakni SIBI dan BISINDO [5]. Namun, karena BISINDO dibuat secara langsung oleh kelompok Tuli, BISINDO menjadi bahasa isyarat yang paling umum digunakan di Indonesia [6].

Kecerdasan Buatan (AI) mencakup berbagai teknik yang memungkinkan komputer meniru perilaku manusia dan mampu memecahkan tugas-tugas kompleks secara mandiri atau dengan sedikit campur tangan manusia [7]. *Computer Vision* dapat secara signifikan meningkatkan efisiensi, keberulangan, dan akurasi dalam tinjauan gambar. *Computer vision* menggunakan fitur-fitur gambar seperti warna, bentuk, dan tekstur untuk menyimpulkan konten gambar [8]. Fokusnya adalah mengajarkan komputer untuk memahami dan menginterpretasi gambar atau video, memungkinkannya untuk mempersepsi dan menganalisis informasi visual secara mirip dengan manusia [9], [10].

Pada kecerdasan buatan atau AI terdapat *Machine learning* yang merupakan bidang studi dalam ilmu komputer dan kecerdasan buatan yang berfokus pada pengembangan algoritma dan model yang mampu belajar secara otomatis dan membuat prediksi atau keputusan tanpa diprogram secara eksplisit. Sementara itu, pada *Computer Vision*, *Machine Learning* (ML) memiliki peran penting dalam mengekstrak informasi penting dari gambar. *Computer vision* menggabungkan metode dari ilmu komputer, pengolahan citra, dan machine learning untuk mengekstraksi informasi yang bermakna dari data visual [9]. Algoritma *machine learning* dilatih menggunakan data dan dapat mempelajari pola dan hubungan dalam data tersebut untuk melakukan prediksi atau tugas tertentu [7].

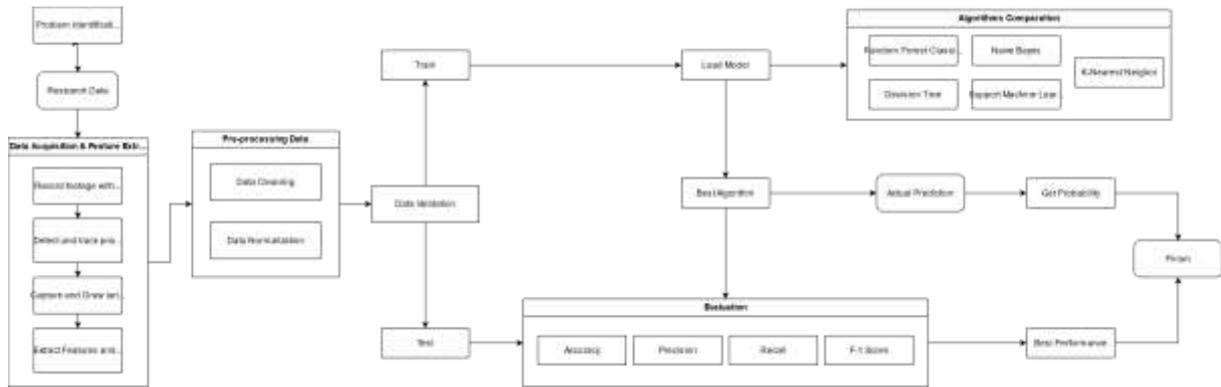
Telah banyak dilakukan pendekatan menggunakan kamera dan *Computer Vision* dalam

menginterpretasikan bahasa isyarat. Namun, jika kita membahas mengenai identifikasi dan pengenalan postur, gerakan, proksemik, dan perilaku manusia, maka subjek yang lebih sesuai adalah *Gesture Recognition*. *Gesture recognition* dapat dilihat sebagai cara komputer untuk memahami bahasa tubuh manusia [11]. *Gesture Recognition* sendiri merupakan topik dalam ilmu komputer dan teknologi bahasa dengan tujuan menafsirkan gerakan manusia melalui algoritma matematika [12]. Dengan sistem *gesture recognition* ini, sudah bukan tidak mungkin untuk mengembangkan sistem yang bisa menerjemahkan bahasa isyarat.

Untuk keperluan *Gesture Recognition* atau deteksi objek, terdapat beberapa framework yang dapat digunakan untuk mempermudah *tracking object*, framework tersebut adalah *MediaPipe*. *MediaPipe* merupakan sebuah framework lintas platform yang bersifat *open-source* yang dikembangkan oleh Google. Framework ini memberikan infrastruktur pipa yang serbaguna untuk memproses data perseptual, seperti video dan audio, secara *real-time*. *MediaPipe* menawarkan berbagai komponen dan algoritma yang sudah siap digunakan untuk tugas-tugas seperti deteksi objek, deteksi wajah, estimasi postur, dan lainnya [13]. Pada kasus deteksi bahasa isyarat, *Mediapipe* digunakan untuk kebutuhan *pose landmark detection* (atau dikenal juga dengan *pose estimation*), *hands landmark detection*, dan *face landmark detection*.

Random Forest telah terbukti bermanfaat dalam banyak bidang karena akurasi prediksi yang tinggi [14], [15]. Di bidang biologi dan kedokteran, *Random Forest* berhasil mengatasi berbagai masalah, termasuk memprediksi respons obat pada sel kanker atau mengidentifikasi protein pengikat DNA. *Random Forest* juga dapat mengenali ucapan dan angka tulisan tangan dengan akurasi tinggi [14]. Algoritma ini dikembangkan oleh Leo Breiman dan umum digunakan untuk tugas klasifikasi dan regresi [16]. Algoritma ini termasuk ke dalam jenis algoritma klasifikasi yang terdiri dari beberapa pohon keputusan dimana setiap pohon keputusan dalam *Random Forest* dibangun berdasarkan nilai vektor sampel yang dipilih secara acak, yang didistribusikan secara independen dan identik untuk setiap pohon [17].

Tujuan penelitian ini adalah mengembangkan sistem penerjemah bahasa isyarat BISINDO dengan fokus pada isyarat ungkapan dan kata-kata dalam bahasa isyarat yang dinilai cukup kompleks. Penelitian bermaksud mengatasi kesenjangan penting dalam studi yang ada, di mana jarang membahas konteks penting ini, meskipun merupakan komponen krusial dari komunikasi bahasa isyarat. Untuk mencapai tujuan tersebut, teknik *Machine Learning* dan *Computer Vision* akan dimanfaatkan, seperti penggunaan *MediaPipe* untuk membantu dalam deteksi objek dan gestur dan penggunaan algoritma *Random Forest* yang akan menjadi algoritma utama karena keandalan dan kemampuannya dalam mengatasi



Gambar 4. Koordinat Face Landmark

data berdimensi tinggi. Selain itu, algoritma populer seperti *K-Nearest Neighbors*, *Naïve Bayes*, *Decision Trees*, dan *Support Vector Machines* juga akan digunakan sebagai pembandingan. Tentunya, kompleksitas penggunaan data kata-kata dalam bahasa isyarat akan mempengaruhi proses pengumpulan dataset dan penentuan model yang paling efektif. Dengan pendekatan ini, diharapkan dapat menciptakan sistem yang tangguh dan akurat untuk menerjemahkan bahasa isyarat, sehingga meningkatkan aksesibilitas dan inklusivitas bagi komunitas tunarungu dan tunawicara.

2. Metode Penelitian

Rancangan penelitian terdiri dari beberapa tahapan, yaitu *data acquisition* (pengumpulan data), *pre-processing data* (pra-pemrosesan data), *modeling* (pemodelan), dan *evaluation* (evaluasi). Rancangan penelitian dapat dilihat pada gambar 1 diatas.

2.1. Identifikasi Masalah

Tantangan atau permasalahan utama dalam penelitian ini meliputi kompleksitas data yang berbasis isyarat kosakata dan penentuan model terbaik. Dalam konteks deteksi bahasa isyarat BISINDO berbasis kosakata, kompleksitas data menjadi faktor krusial yang perlu diatasi. Untuk mengatasi kompleksitas data berbasis isyarat kosakata, diperlukan pendekatan yang tepat dan memakan waktu yang signifikan. Selain itu, penentuan model terbaik yang efektif untuk menginterpretasikan gerakan bahasa isyarat juga menjadi tantangan yang perlu dipecahkan. Dalam upaya mencapai tujuan tersebut, dalam penelitian ini akan diuji beberapa model seperti *Random Forest*, *Support Vector Machines (SVM)*, *Decision Trees*, *K-Nearest Neighbors (KNN)*, dan *Naive Bayes* untuk menentukan model yang paling efektif.

2.2. Data Penelitian

Dalam penelitian ini, Data yang dipakai berbentuk CSV (*Comma Separated Value*) yang dipakai terdiri dari 10 kelas atau kata-kata dalam bahasa isyarat. Setiap kelas memiliki 2500 sampel data. Label data ditentukan berdasarkan nama gestur bahasa isyarat yang mewakili setiap kelas. Seperti yang bisa dilihat pada Tabel 1, fitur yang diambil dari setiap sampel data meliputi koordinat

x, y, dan z, serta v (visibilitas). Koordinat x dan y mencerminkan posisi *landmark* dalam koordinat 2D, sedangkan koordinat z menggambarkan kedalaman *landmark* dalam ruang 3D. Koordinat v atau visibilitas mengindikasikan apakah *landmark* terlihat jelas atau tidak [18].

Tabel 1. Sampel Label & Feature Data Penelitian

Class	x1	y1	z1	...	v501
Apa	0.499...	0.287...	-0.00...	...	0.0
Baik	0.478...	0.296...	-1.026...	...	0.0
Dimana	0.494...	0.200...	-0.671...	...	0.0
Halo	0.571...	0.280...	-0.513...	...	0.0
Kabar	0.551...	0.313...	-0.922...	...	0.0
Kamu	0.534...	0.317...	-0.775...	...	0.0
Maaf	0.442...	0.255...	-0.709...	...	0.0
Saya	0.373...	-0.499...	-0.654...	...	0.0
Teman	0.525...	0.254...	-0.929...	...	0.0
Terima kasih	0.476...	0.311...	-0.659...	...	0.0

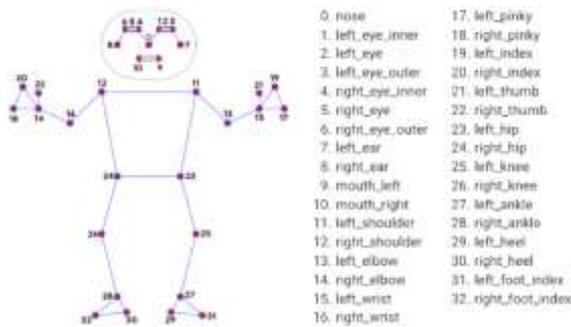
2.3. Data Acquisition

Metode pengumpulan data yang digunakan oleh penulis memiliki tujuan untuk mengumpulkan data yang akurat dan terstruktur mengenai bahasa isyarat. Secara keseluruhan, proses akuisisi data ini terdiri dari empat langkah utama. Pertama, dilakukan perekaman video menggunakan *webcam* secara *real-time*. Kemudian, dengan menggunakan model *MediaPipe*, dilakukan deteksi dan pelacakan landmark pada pose, tangan, dan wajah yang terdapat dalam setiap *frame* video. Langkah berikutnya adalah menampilkan *landmark* yang telah dideteksi tersebut pada video yang direkam. Terakhir, dilakukan ekstraksi fitur dari *landmark* yang berhasil ditangkap dan hasilnya diekspor ke dalam format CSV (*Comma-Separated Values*), yang memudahkan analisis dan pengolahan lebih lanjut.

Penggunaan *MediaPipe* pada langkah kedua dan ketiga sebagai alat untuk ekstraksi *landmark* memberikan keunggulan dalam mengumpulkan data secara objektif dan terstandarisasi. *Mediapipe* juga secara akurat memperkirakan koordinat 2D atau 3D dari titik-titik. Dengan menggunakan koordinat *landmark* sebagai fitur utama, data yang terkumpul dapat memberikan informasi yang berguna untuk analisis lebih lanjut [18],

[19], seperti pengembangan model pengenalan bahasa isyarat atau penelitian lebih mendalam mengenai gerakan tubuh manusia.

Pose landmark detection digunakan untuk mendeteksi landmark tubuh manusia [18]. *MediaPipe Holistic* akan mendeteksi pose/sikap tubuh seseorang dalam frame video. Hal ini melibatkan penentuan lokasi komponen penting tubuh seperti kepala, bahu, lengan, pinggul, dan kaki yang memungkinkan aplikasi seperti pelacakan gerakan, dan pengenalan gerakan. *MediaPipe Holistic* melakukan tugas ini menggunakan model *machine learning* yang bernama *BlazePose* [18], [20]. Pada penelitian ini, pengambilan data kebanyakan hanya mengambil titik koordinat dari kepala hingga bahu atau hingga pinggang. Gambar 2 dibawah menunjukkan detail atribut koordinat *pose landmark* yang telah disediakan oleh *MediaPipe*.



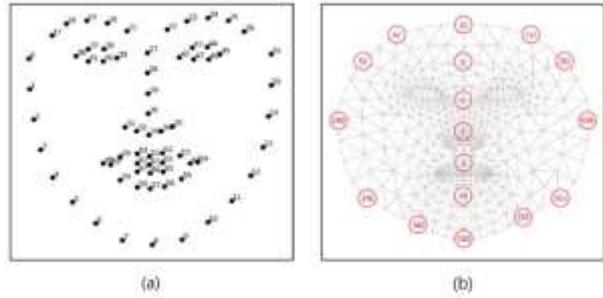
Gambar 2. Koordinat Pose Landmark

Hand landmark MediaPipe berfokus pada lokalisasi titik-titik penting tangan manusia, termasuk telapak tangan, jari-jari, dan sendi - sendi jari. seperti *pose landmark*. Model *hand landmark* ini mendeteksi 21 keypoint atau titik koordinat di masing-masing tangan [18], [20]. Gambar dibawah menunjukkan detail atribut koordinat *Hand Landmark* yang telah disediakan oleh *MediaPipe*.



Gambar 3. Koordinat Hand Landmark

Face landmark mendeteksi kurang lebih 468 keypoints pada wajah. *Face landmark* juga melibatkan deteksi dan pelacakan titik-titik penting pada wajah manusia, seperti mata, alis, hidung, mulut, dan garis rahang. Model ini dilatih untuk mengatasi berbagai ekspresi dari wajah, pose kepala, maupun kondisi pencahayaan yang berbeda [18], [20]. Keandalan fungsi ini memungkinkannya untuk mengenali dan melacak titik-titik tersebut secara akurat pada berbagai kondisi. Berikut detail persebaran dari koordinat *Face Landmark* yang telah disediakan oleh *MediaPipe* ditunjukkan pada gambar 4.



Gambar 4. Koordinat Face Landmark

2.4. Pre-processing Data

Tahap *pre-processing data* sangat penting dalam mempersiapkan data untuk analisis atau pemodelan selanjutnya. Proses ini terutama melibatkan *data cleaning* dan *data normalization* untuk memastikan kualitas dan konsistensi data. Selama *pre-processing data*, salah satu tugas utama adalah mengidentifikasi dan menangani *missing value* pada data [21]. Data di penelitian ini memiliki format CSV, maka jika ada baris atau kolom yang mempunyai missing value, baris tersebut akan dihapus dari kumpulan data, memastikan hanya baris yang lengkap dan dapat diandalkan yang tetap dipertahankan. *Pre-processing data* juga melibatkan identifikasi *anomali data* yang tidak memenuhi standar yang diperlukan atau dianggap terlalu tidak akurat atau tidak dapat diandalkan. *Data anomali* ini, yang dapat mempengaruhi proses analisis atau pemodelan, dapat diidentifikasi dan dihapus dari kumpulan data. Dengan menghilangkan data yang hilang dan data anomali, kumpulan data yang telah dibersihkan menjadi lebih cocok untuk analisis selanjutnya.

2.5. Data Validation

Pada tahap *data validation*, metode yang digunakan adalah *Split Validation*. Pendekatan ini penting untuk mendapatkan evaluasi yang tidak bias terhadap performa prediksi model dan memvalidasi model tersebut [22], [23]. Pada *Split Validation*, dataset yang telah melalui proses pembersihan data akan dibagi menjadi dua bagian, yaitu data latih (*train*) dan data uji (*test*). Data latih digunakan untuk melatih model, memungkinkannya untuk mempelajari pola dan hubungan dalam data [23]. Sementara itu, data uji digunakan untuk mengevaluasi kinerja model yang telah dilatih pada data yang tidak terlihat sebelumnya, sehingga memberikan perkiraan sejauh mana model akan menggeneralisasi pada contoh-contoh baru [22]. Pembagian data ini dilakukan sebelum melakukan proses apa pun pada data, baik itu melatih model *Machine Learning* maupun melakukan analisis statistik untuk mengidentifikasi fitur-fitur yang menarik. Umumnya, pembagian data menggunakan perbandingan 80% untuk data latih dan 20% untuk data uji, namun perbandingan tersebut dapat disesuaikan sesuai dengan kebutuhan. Dengan menggunakan metode *Split Validation*, evaluasi kinerja model akan menjadi lebih

obyektif karena menggunakan data yang belum pernah dilihat oleh model sebelumnya [22]. Hal ini memastikan bahwa performa prediksi model dievaluasi dengan data yang baru dan tidak dipengaruhi oleh data latih [23].

2.6. Komparasi Algoritma

Dalam proses pemodelan, setelah data dibagi menjadi data pelatihan dan data uji, maka data tersebut akan digunakan untuk melatih dan menguji performa masing-masing algoritma. Tujuannya adalah memilih algoritma yang paling efektif dalam mengenali dan menginterpretasikan gestur bahasa isyarat secara akurat. Penelitian ini menguji sebanyak lima algoritma berbeda.

Random Forest Classifier merupakan salah satu algoritma yang termasuk dalam kelompok *supervised learning* yang dikembangkan oleh Leo Breinman. Algoritma ini umumnya digunakan untuk tugas klasifikasi dan regresi. *Random Forest* merupakan suatu algoritma klasifikasi yang terdiri dari beberapa pohon keputusan. Setiap pohon keputusan dalam *Random Forest* dibangun berdasarkan nilai vektor sampel yang dipilih secara acak, yang didistribusikan secara independen dan identik untuk setiap pohon [17].

Algoritma *K-Nearest neighbour* (KNN) merupakan algoritma *machine learning* yang terutama digunakan untuk tujuan klasifikasi. Secara umum, algoritma KNN dapat mengklasifikasikan dataset dengan menggunakan model pelatihan yang serupa dengan *query* pengujian dengan mempertimbangkan k data pelatihan terdekat (tetangga). Kemudian, algoritma ini melakukan aturan pemungutan suara mayoritas untuk menentukan klasifikasi akhir. Di antara semua algoritma *machine learning*, algoritma KNN merupakan salah satu yang paling sederhana dan banyak digunakan dalam subjek klasifikasi karena desainnya yang sangat adaptif dan mudah dipahami [24].

Algoritma *Naive Bayes* adalah algoritma klasifikasi probabilitas sederhana yang menghitung serangkaian probabilitas dengan menghitung frekuensi dan kombinasi nilai-nilai dalam sebuah dataset yang diberikan. Algoritma ini menggunakan *teorema Bayes* dan mengasumsikan bahwa semua variabel adalah independen dengan mempertimbangkan nilai dari variabel kelas. Asumsi independensi kondisional ini jarang valid dalam aplikasi dunia nyata, sehingga algoritma ini disebut sebagai *Naive* (sederhana), namun algoritma ini cenderung belajar dengan cepat dalam berbagai masalah klasifikasi yang terkontrol [25].

Decision Tree merupakan salah satu metode yang umum digunakan dalam *machine learning*, pengolahan citra, dan identifikasi pola [26]. Sebuah pohon biasa terdiri dari akar (*root*), cabang (*branches*), dan daun (*leaves*). Struktur yang sama ditiru oleh *Decision Tree*. Algoritma ini terdiri dari simpul akar (*root node*), cabang-cabang, dan simpul daun (*leaf nodes*). *Decision Tree* bisa dianggap sebagai sebuah pohon di mana setiap simpul mewakili suatu fitur (atribut), setiap tautan (cabang)

mewakili suatu keputusan (aturan), dan setiap daun mewakili hasil (nilai kategori atau nilai kontinu) [27].

Support Vector Machine (SVM) memiliki keunggulan dalam hal teori yang lengkap, optimasi global, adaptabilitas yang kuat, dan kemampuan generalisasi yang baik karena didasarkan pada Teori Pembelajaran Statistik (*Statistical Learning Theory/SLT*) [28]. Dalam model SVM, sebuah *hyperplane* keputusan dibangun untuk membentuk celah pemisahan guna membagi dua contoh kelas dengan margin maksimum [29]. SVM memiliki keunggulan unik dalam menangani pengenalan pola nonlinear, dan dimensi yang tinggi [28].

Setelah melatih model dengan menggunakan data pelatihan, tahap pengujian dilakukan menggunakan data uji yang sebelumnya telah dipisah. Pada tahap ini, performa masing-masing algoritma akan dievaluasi dengan membandingkan hasil prediksi mereka dengan label yang sebenarnya.

2.7. Evaluasi

Di tahap evaluasi, peneliti menggunakan metrik evaluasi seperti skor akurasi (*accuracy score*) yang digunakan untuk mengevaluasi kinerja dari model yang diujikan dan juga laporan klasifikasi (*classification report*) yang digunakan memberikan informasi lebih rinci tentang kinerja model dalam mengenali setiap kelas atau label gestur bahasa isyarat. Metrik yang umum digunakan dalam laporan klasifikasi adalah *recall*, *precision*, dan *F1-Score* [30].

Akurasi (*accuracy*) merupakan salah satu metrik evaluasi yang digunakan untuk mengukur sejauh mana model *machine learning* mampu mengklasifikasikan data dengan benar [31]. Dalam konteks pengenalan gestur bahasa isyarat, akurasi mengindikasikan seberapa baik model dapat mengenali dan memprediksi gestur bahasa isyarat dengan benar. Berikut adalah rumus (1) dari nilai Akurasi [32].

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

Accuracy dihitung dengan cara menghitung hasil pembagian dari jumlah prediksi benar, $TP + TN$ (*True Positive + True Negative*) dengan besar dataset, $TP + TN + FP + FN$ (*True Positive + True Negative + False Positive + False Negative*). *Recall* (sensitivitas) mengukur sejauh mana model mampu mengidentifikasi dengan benar gestur bahasa isyarat dari suatu kelas. *Recall* dapat dianggap sebagai kemampuan model dalam menemukan semua contoh positif dari suatu kelas. Berikut adalah formula (2) dari recall.

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

Recall dihitung dengan membagi jumlah TP (*True Positive*) atau prediksi benar positif dengan jumlah total data aktual yang sebenarnya positif yakni $TP + FN$ (*True Positive + False Negative*) [31]. *Precision* atau presisi mengindikasikan seberapa akurat model dalam

mengenali gestur bahasa isyarat dari suatu kelas dan menghindari memberikan hasil yang salah positif. Berikut merupakan formula atau rumus (3) dari *precision*. *Precision* dihitung dengan membagi jumlah *TP* (*True Positive*) prediksi benar positif dengan jumlah total prediksi positif, *TP + FP* (*True Positive + False Positive*) [31].

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

Terakhir adalah *F1-Score* yang merupakan perpaduan antara *recall* dan *precision*. *F1-Score* memberikan gambaran keseluruhan tentang kinerja model dengan mempertimbangkan kedua metrik tersebut. *F1-Score* yang tinggi menunjukkan keseimbangan yang baik antara *recall* dan *precision* [37], [38]. *F1 Score* dihitung dengan mengambil nilai rata-rata harmonik dari *precision* dan *recall*. Berikut merupakan formula (4) dari *F1-Score*. Rumus tersebut menghitung rata-rata harmonik dari nilai *precision* dan *recall* [31].

$$F1\ Score = \frac{2}{\frac{1}{precision} + \frac{1}{Recall}} = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

Secara keseluruhan, tahap evaluasi studi ini menggunakan berbagai metrik evaluasi untuk menilai kinerja model yang diuji dalam mengenali dan memprediksi gestur bahasa isyarat. Metrik evaluasi yang digunakan meliputi skor akurasi, *recall*, *precision*, dan *F1-score*. Dengan menggunakan metrik-metrik evaluasi ini, peneliti memperoleh wawasan berharga tentang kinerja model dan kemampuannya dalam mengenali dan mengklasifikasikan gestur bahasa isyarat secara akurat. Temuan dari tahap evaluasi ini menjadi dasar analisis penelitian ini dan memberikan kontribusi pada pemahaman mengenai efektivitas model dalam konteks deteksi atau rekognisi gestur bahasa isyarat.

2.8. Actual Prediction

Tahap prediksi pada penelitian ini mengacu pada output yang diperoleh dari model setelah menganalisis gerakan isyarat yang diberikan, yang menunjukkan kelas atau label yang ditafsirkan. Selain itu, sistem *gesture recognition* sering memberikan probabilitas deteksi, yang mewakili tingkat kepercayaan yang terkait dengan prediksi. Probabilitas ini menunjukkan tingkat kepastian model dalam prediksinya, dengan probabilitas yang lebih tinggi menunjukkan kepercayaan yang lebih besar. Melacak probabilitas deteksi memungkinkan untuk menilai keandalan prediksi dan menetapkan ambang batas yang sesuai.

3. Hasil dan Pembahasan

Pada Penelitian ini, seperti yang dibahas sebelumnya, data yang digunakan merupakan data yang diambil secara manual berisikan nilai koordinat dari gestur tubuh (pose, gerakan tangan, wajah) saat sedang menggunakan bahasa isyarat. Data ini memuat label yang berisi kata-

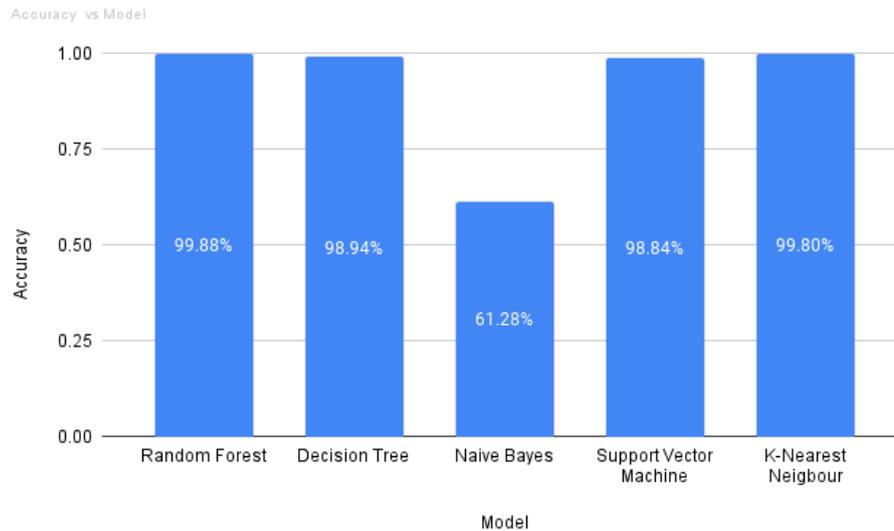
kata bahasa Indonesia dalam bahasa isyarat BISINDO, terdiri dari 10 kata, yakni: Apa, Baik, Dimana, Halo, Kabar, Kamu, Maaf, Saya, Teman, Terima kasih. Kata-kata tersebut dipilih berdasarkan variasi dan kompleksitasnya. Ada kata yang menggunakan gerakan satu tangan, dan gerakan dua tangan, ada juga kata yang memiliki pose gerakan yang cukup mirip seperti gerakan kata maaf dengan kabar dan kata kamu dengan baik. Gambar 6 di bawah menunjukkan sampel dari gestur bahasa isyarat yang dipakai sebagai data.



Gambar 6. Sampel Kosa Kata

Setiap label atau kata memiliki 2500 sampel data, yang mana secara keseluruhan data ini berjumlah 25.000 sampel. Fungsi dari *framework mediapipe* digunakan untuk melakukan deteksi *landmark* pada *frame* saat sedang melakukan koleksi data dengan perekaman secara real-time. *Mediapipe* juga digunakan untuk menggambar *landmark* pada *frame*, yakni *landmark* wajah, pose tubuh, tangan kiri, dan tangan kanan pada gambar. Nilai koordinat *landmark* ini yang diambil dan disimpan sebagai data. Jumlah koordinat *landmark* yang akan disimpan dihitung berdasarkan jumlah *landmark* pose tubuh dan wajah. Koordinat ini terdiri dari 'x' dan 'y' untuk koordinat titik *landmark*, 'z' untuk koordinat kedalaman, dan 'v' untuk visibilitas (*visibility*) *landmark*. Ini berfungsi sebagai faktor kunci dalam menentukan dimensi dan organisasi dataset, sehingga memungkinkan analisis dan pengenalan gestur bahasa isyarat yang efektif. Pada penelitian ini didapat sebanyak 501 koordinat *landmark* yang akan dianalisis dan diekstraksi untuk mempelajari dan mengenali gestur bahasa isyarat. Data tersebut kemudian disimpan dalam format CSV (*Comma Separated Value*).

Data yang sudah didapatkan tidak langsung diimplementasikan model algoritma, melainkan melalui beberapa proses yakni *data cleaning* dan *normalization*, yang termasuk ke dalam tahap *pre-processing data*. Pada proses ini, data yang memiliki *missing value* akan langsung di *drop*/dihapus, akan tetapi karena proses pengambilannya yang manual, masih dilakukan juga pemeriksaan secara manual untuk memastikan format data sudah benar atau belum. *Split-Validation* kemudian dilakukan untuk membagi data menjadi data latih (*train*) dan data uji (*test*). Perbandingan data latih dan uji di penelitian kali ini yaitu 80:20, dimana data latih dengan 80% bagian dan data uji memiliki 20% sisanya.



Gambar 7. Grafik Komparasi Algoritma

Pengujian model pada penelitian ini dilakukan pada 5 algoritma yakni: *Random Forest Classifier*, *Decision Tree*, *Support Vector Machine*, *K-Nearest Neighbor* dan *Naive Bayes*. Tabel 2 memuat *Classification Report* yang berisi hasil evaluasi performa model – model tersebut. Performa model diukur dengan menggunakan metrik *precision* (presisi), *recall* (ingat), *F1-Score*, dan akurasi. Tabel tersebut menunjukkan bahwa *Random Forest* dan *SVM* merupakan model terbaik dalam deteksi bahasa isyarat dengan performa yang sangat tinggi. Sedangkan *Decision Tree* dan *KNN* juga memberikan hasil yang baik, meskipun sedikit lebih rendah dibandingkan dengan dua model sebelumnya.

Tabel 2. Classification Report 5 Model

Model	Precision	Recall	F1-Score	Accuracy
Random Forest	99.88%	99.88%	99.88%	99.88%
Support Vector Machine	99.80%	99.80%	99.80%	99.80%
Decision Tree	98.94%	98.94%	98.94%	98.94%
K-Nearest Neighbour	98.84%	98.84%	98.84%	98.84%
Naive Bayes	62.90%	61.28%	61.01%	61.28%

Sementara itu untuk komparasi dari rata-rata akurasi dari semua model algoritma ditunjukkan pada grafik di gambar 7. Pada grafik tersebut, dapat dilihat bahwa model *Random Forest* memiliki nilai akurasi tertinggi dengan persentase 99.88% sedangkan nilai akurasi terendah dipegang oleh model *Naive Bayes* dengan persentase 61.28%. Setelah membandingkan hasil evaluasi yang telah dilakukan, algoritma *Random Forest* terpilih sebagai algoritma terbaik karena memberikan kinerja yang paling baik dalam memprediksi kelas-kelas pada dataset yang digunakan.

Terdapat atribut-atribut penting yang menunjukkan kinerja model deteksi bahasa isyarat menggunakan algoritma *Random Forest*, Hal ini dapat dilihat pada

Tabel 3. Tabel tersebut memuat hasil *classification report* untuk algoritma *Random Forest*. *Precision* menggambarkan kemampuan model dalam mengidentifikasi dengan benar label positif. Nilai *precision* yang tinggi, seperti 100.00% untuk kata "Apa," menunjukkan akurasi yang sangat baik dalam mengklasifikasikan setiap kata bahasa isyarat secara tepat. *Recall* menggambarkan kemampuan model dalam menemukan kembali (mengingat) label positif secara keseluruhan. *Recall* yang tinggi, seperti 100.00% untuk kata "Baik," menunjukkan kemampuan yang sempurna dalam mendeteksi gerakan isyarat yang relevan untuk kata tersebut. *F1 Score* merupakan ukuran gabungan antara *precision* dan *recall* yang memberikan gambaran keseimbangan antara kemampuan mengklasifikasikan dengan benar dan kemampuan menemukan kembali gerakan isyarat yang relevan. *F1 Score* yang tinggi, mencapai 99.90% atau 100.00% untuk setiap kata dalam tabel, menunjukkan performa yang sangat baik secara keseluruhan. *Support* adalah nilai yang mencerminkan jumlah sampel dalam set data yang mendukung setiap kata. *Support* yang lebih tinggi menunjukkan adanya lebih banyak contoh gerakan isyarat yang relevan dalam dataset. Terakhir, akurasi (*Accuracy*) menggambarkan sejauh mana model secara keseluruhan dapat mengklasifikasikan dengan benar seluruh dataset. Pada penelitian ini, akurasi mencapai 99.88%, menunjukkan tingkat keberhasilan yang sangat tinggi dalam mengenali gerakan isyarat dan mengklasifikasikan kata dalam bahasa isyarat dengan benar. Selain itu, terdapat atribut "macro avg" dan "weighted avg" yang memberikan rata-rata *precision*, *recall*, dan *F1 Score* dari seluruh kelas kata dalam dataset. *Macro Average* adalah rata-rata metrik untuk setiap kelas, sedangkan *Weighted Average* adalah rata-rata yang memperhitungkan jumlah data dalam setiap kelas. Rata-rata dari atribut-atribut ini dalam tabel mencapai 99.88%, menunjukkan kinerja yang sangat baik secara keseluruhan dalam deteksi bahasa isyarat menggunakan algoritma *Random Forest*.

Tabel 3. Random Forest Classification Report

No	Kata	Random Forest			Support
		Precision	Recall	F1 Score	
1	Apa	100.00%	99.80%	99.90%	493
2	Baik	99.81%	100.00%	99.90%	518
3	Dimana	100.00%	100.00%	100.00%	515
4	Halo	100.00%	100.00%	100.00%	504
5	Kabar	99.36%	99.79%	99.57%	466
6	Kamu	100.00%	100.00%	100.00%	512
7	Maaf	99.80%	99.39%	99.59%	494
8	Saya	100.00%	100.00%	100.00%	507
9	Teman	99.79%	99.79%	99.79%	482
10	Terima Kasih	100.00%	100.00%	100.00%	511
11	Accuracy			99.88%	5,002
12	Macro avg	99.88%	99.88%	99.88%	5002
13	Weighted avg	99.88%	99.88%	99.88%	5002

Melalui hasil uji yang telah dilakukan, algoritma *Random Forest* telah menghasilkan hasil yang sangat baik dalam klasifikasi pada setiap kata/label. Pada beberapa kata seperti "Apa," "Baik," "Dimana," "Halo," "Kamu," "Saya," "Teman," dan "Terima Kasih," algoritma ini mampu mencapai tingkat *precision*, *recall*, dan *F1-score* yang sempurna atau mendekati sempurna, dengan nilai 100% pada masing-masing metrik evaluasi. Ini menunjukkan bahwa algoritma *Random Forest* sangat efektif dalam mengklasifikasikan kata-kata tersebut. Meskipun ada beberapa kata seperti "Kabar" dan "Maaf" yang tidak mencapai tingkat sempurna, namun hasil yang didapat masih sangat tinggi, dengan persentase *precision*, *recall*, dan *F1-score* mencapai 99%. Hasil evaluasi ini menegaskan bahwa *Random Forest* adalah algoritma terbaik untuk tugas klasifikasi gestur kosakata dalam bahasa isyarat BISINDO yang diberikan.

Selain melakukan evaluasi pada klasifikasi kata/label, penelitian ini juga mencoba untuk membuktikan apakah sistem yang dikembangkan dapat secara efektif mengenali gerakan isyarat (*gesture recognition*) dengan menggunakan model terbaik yang telah dipilih, yaitu model *random forest classifier*. Namun, meskipun telah memiliki model terbaik, terkadang sistem masih mengalami kesulitan dalam memprediksi gerakan secara langsung. Untuk menguji kebenaran prediksi secara aktual, penulis melakukan pengujian dengan menghitung tingkat probabilitas deteksi pada berbagai gerakan isyarat.

Seperti yang bisa dilihat pada Tabel 4, hasil actual prediction dalam penelitian ini dapat dikatakan cukup bagus. Dengan rata-rata probabilitas kebenaran prediksi antara 0,50 hingga 0,60, sistem prediksi dianggap sudah cukup baik dan dapat digunakan. Terutama pada kata "Halo" yang memiliki probabilitas deteksi tinggi sebesar 0,71, sistem mampu memprediksi dengan sangat cepat

dan efisien. Namun, terdapat beberapa kata seperti "Dimana" dengan probabilitas di bawah 0,30, di mana terkadang kata tersebut masih sering tertukar dengan kata-kata lain yang memiliki pose dan gerakan yang mirip, seperti "Maaf" dan "Terima Kasih". Hal ini menunjukkan bahwa masih ada beberapa tantangan dalam pengenalan gerakan isyarat, terutama ketika gerakan memiliki kemiripan secara visual.

Tabel 4. Sampel Actual Prediction Report

Frame	Probabilitas Deteksi	Isyarat
	0.71	Halo
	0.52	Apa
	0.38	Maaf
	0.25	Dimana

Secara keseluruhan, penelitian ini memberikan temuan hasil yang lebih relevan dan inovatif dalam sistem deteksi bahasa isyarat dengan memperluas cakupan deteksi kata-kata dalam bahasa isyarat BISINDO. Sebelumnya, penelitian dengan bidang atau konteks serupa sering kali hanya memusatkan perhatian pada gestur abjad atau angka dalam bahasa isyarat. Namun, dengan memasukkan tipe kosakata dalam proses deteksi, penelitian ini menghadirkan pengenalan bahasa isyarat yang lebih realistis dan sesuai dengan penggunaan sehari-hari dalam komunikasi. Dalam hal ini algoritma *Random Forest* juga berhasil diimplementasikan pada penelitian dan bahkan mencapai tingkat akurasi yang sangat tinggi, yaitu 99,88%. Tingkat akurasi yang mencapai angka sedemikian tinggi ini menjadi pencapaian yang menonjol dan menandai kesuksesan pendekatan yang telah diusulkan. Penemuan ini memberikan pemahaman yang lebih komprehensif tentang bahasa isyarat dan membuka peluang baru untuk pengembangan lebih lanjut dalam bidang ini.

4. Kesimpulan

Penelitian ini telah berhasil memanfaatkan *Mediapipe* dan algoritma *Random Forest Classifier* sebagai alat bantu untuk pembuatan sistem deteksi bahasa isyarat dengan data label kosakata yang memiliki kompleksitas yang tinggi. Hasil evaluasi menunjukkan bahwa

algoritma *Random Forest* mengungguli model lainnya, dengan mencapai *precision*, *recall*, *F1-score*, dan akurasi keseluruhan tertinggi (99,88%). Hal ini menekankan efektivitas *Random Forest* dalam mengklasifikasikan gerakan isyarat dan kata secara akurat. Laporan klasifikasi juga mengkonfirmasi kinerja luar biasa dari model *Random Forest*, dengan mencapai *presisi*, *recall*, dan *F1-score* yang tinggi untuk sebagian besar label, bahkan mencapai 100% pada beberapa label. Hal ini menunjukkan kemampuan model dalam mengenali dan memprediksi kata-kata tertentu dengan akurat. Namun, dengan kompleksitas data yang ada terdapat tantangan dalam membedakan gerakan isyarat yang mirip, yang mengakibatkan kesalahan klasifikasi terkadang dapat terjadi. Meskipun demikian, kinerja keseluruhan sistem dalam sistem deteksi bahasa isyarat dapat dikatakan baik, dengan probabilitas prediksi yang berkisar antara 0,50 hingga 0,70 untuk prediksi yang benar. Secara keseluruhan, penelitian ini memberikan hasil yang signifikan dalam pengembangan sistem pengenalan gerakan isyarat. Dengan memilih algoritma *Random Forest* sebagai model terbaik, penelitian ini memberikan landasan yang kuat untuk pengembangan lebih lanjut dalam teknologi pengenalan gerakan isyarat.

Daftar Pustaka

- [1] World Health Organization, "Deafness and hearing loss," *World Health Organization*, 2023. <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss> (accessed Jun. 04, 2023).
- [2] S. Rabiah, "Language as a tool for communication and cultural reality discloser," 2018.
- [3] L. R. Nendauni, *The Development Of Sign Language: A Synopsis Overview*. 2021. doi: 10.13140/RG.2.2.19207.93609.
- [4] A. Halder and A. Tayade, "Real-time vernacular sign language recognition using mediapipe and machine learning," *Journal homepage: www.ijrpr.com ISSN*, vol. 2582, p. 7421, 2021.
- [5] N. P. L. Wedayanti, "Teman Tuli diantara SIBI dan BISINDO," in *Proceedings*, 2019, pp. 137–146.
- [6] Klobility, "BISINDO dan SIBI: Apa Bedanya?," *Klobility*, 2019. <https://www.klobility.id/post/perbedaan-bisindo-dan-sibi> (accessed Jun. 04, 2023).
- [7] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electronic Markets*, vol. 31, no. 3, pp. 685–695, 2021, doi: 10.1007/s12525-021-00475-2.
- [8] B. G. Weinstein, "A computer vision for animal ecology," *Journal of Animal Ecology*, vol. 87, no. 3, pp. 533–545, May 2018, doi: <https://doi.org/10.1111/1365-2656.12780>.
- [9] A. Ayub Khan, A. Laghari, S. Awan, Lyari, and P. Karachi, "Machine Learning in Computer Vision: A Review," *ICST Transactions on Scalable Information Systems*, vol. 8, Apr. 2021, doi: 10.4108/eai.21-4-2021.169418.
- [10] M. Nixon and A. Aguado, *Feature extraction and image processing for computer vision*. Academic press, 2019.
- [11] C. E. Kurian, C. Martin, P. J. Deepthi, and K. A. Eldhose, "Air Writing Recognition and Speech Synthesis," 2018.
- [12] J. Kobylarz, J. J. Bird, D. R. Faria, E. P. Ribeiro, and A. Ekárt, "Thumbs up, thumbs down: non-verbal human-robot interaction through real-time EMG classification via inductive and supervised transductive transfer learning," *J Ambient Intell Humaniz Comput*, vol. 11, no. 12, pp. 6021–6031, 2020, doi: 10.1007/s12652-020-01852-z.
- [13] K. M. Kavana and N. R. Suma, "Recognition of Hand Gestures Using Mediapipe Hands," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 4, no. 06, 2022.
- [14] D. Denisko and M. M. Hoffman, "Classification and interaction in random forests," *Proceedings of the National Academy of Sciences*, vol. 115, no. 8, pp. 1690–1692, 2018.
- [15] J. L. Speiser, M. E. Miller, J. Tooze, and E. Ip, "A comparison of random forest variable selection methods for classification prediction modeling," *Expert Syst Appl*, vol. 134, pp. 93–101, 2019, doi: <https://doi.org/10.1016/j.eswa.2019.05.028>.
- [16] M. Soni and S. Varma, "Diabetes prediction using machine learning techniques," *International Journal of Engineering Research & Technology (Ijert) Volume*, vol. 9, 2020.
- [17] I. Hendapratama, I. W. Hamzah, and S. Astuti, "Rancang Bangun Aplikasi Penerjemah SIBI (Sistem Isyarat Bahasa Indonesia) Menggunakan Algoritma Random Forest Classifier," *eProceedings of Engineering*, vol. 9, no. 6, 2023.
- [18] MediaPipe, "MediaPipe Solutions guide," 2023. <https://developers.google.com/mediapipe/solutions/guide> (accessed Jun. 11, 2023).
- [19] C. Lugaresi *et al.*, "Mediapipe: A framework for building perception pipelines," *arXiv preprint arXiv:1906.08172*, 2019.
- [20] K. Goyal, "Indian Sign Language Recognition Using Mediapipe Holistic," *arXiv preprint arXiv:2304.10256*, 2023.
- [21] Tableau, "Guide To Data Cleaning: Definition, Benefits, Components, And How To Clean Your Data," Oct. 17, 2021. <https://www.tableau.com/learn/articles/what-is-data-cleaning> (accessed Jun. 11, 2023).
- [22] Mirko Stojiljković, "Split Your Dataset With scikit-learn's train_test_split()," *Real Python*, 2023. <https://realpython.com/train-test-split/>

- python-data/#the-importance-of-data-splitting (accessed Jul. 05, 2023).
- [23] G. Varoquaux and O. Colliot, “Evaluating Machine Learning Models and Their Diagnostic Value,” in *Machine Learning for Brain Disorders*, O. Colliot, Ed., New York, NY: Springer US, 2023, pp. 601–630. doi: 10.1007/978-1-0716-3195-9_20.
- [24] S. Uddin, I. Haque, H. Lu, M. A. Moni, and E. Gide, “Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction,” *Sci Rep*, vol. 12, no. 1, p. 6256, 2022, doi: 10.1038/s41598-022-10358-x.
- [25] M. M. Saritas and A. Yasar, “Performance analysis of ANN and Naive Bayes classification algorithm for data classification,” *International journal of intelligent systems and applications in engineering*, vol. 7, no. 2, pp. 88–91, 2019.
- [26] B. Charbuty and A. Abdulazeez, “Classification Based on Decision Tree Algorithm for Machine Learning,” *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.
- [27] H. H. Patel and P. Prajapati, “Study and analysis of decision tree based classification algorithms,” *International Journal of Computer Sciences and Engineering*, vol. 6, no. 10, pp. 74–78, 2018.
- [28] Z. Jun, “The Development and Application of Support Vector Machine,” *J Phys Conf Ser*, vol. 1748, no. 5, p. 052006, 2021, doi: 10.1088/1742-6596/1748/5/052006.
- [29] Y. Ramdhani, R. T. Prasetyo, R. Hidayat, and D. P. Alamsyah, “Comparison of SVM Algorithm and Neural Network With Feature Optimization Based on Genetic Algorithm in Determining Immunotherapy Success in Cancer Disease,” in *Proceedings of the International Conference on Industrial Engineering and Operations Management Monterrey*, 2021, pp. 3142–3149.
- [30] Aman Kharwal, “Classification Report in Machine Learning,” *The Clever Programmer*, Jul. 07, 2021. <https://thecleverprogrammer.com/2021/07/07/classification-report-in-machine-learning/> (accessed Jun. 11, 2023).
- [31] Aaron Zhu, “Essential Evaluation Metrics for Classification Problems in Machine Learning,” *Towards Data Science*, Mar. 10, 2023. <https://towardsdatascience.com/essential-evaluation-metrics-for-classification-problems-in-machine-learning-69e90665375b> (accessed Jun. 11, 2023).
- [32] Teemu Kanstrén, “A Look at Precision, Recall, and F1-Score,” *towards data science*, 2020. <https://towardsdatascience.com/a-look-at-precision-recall-and-f1-score-36b5fd0dd3ec> (accessed Jul. 14, 2023).