Autentikasi Pengguna Berbasis *Keystroke Dynamic* Menggunakan Fitur *Distance Enhanced Flight Time* (DEFT)

1st Mohammad Dwiantara Mahardhika
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
dwiantaramd@student.telkomuniversity.ac.id

2nd Prasti Eko Yunanto
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
gppras@telkomuniversity.ac.id

3rd Febryanti Sthevanie
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
febryantisthevanie@telkomuniversity.ac.id

***Abstrak* —Privasi dan keamanan data menjadi sangat penting di era digital saat ini. Metode autentikasi tradisional yang sering digunakan seperti kata sandi dan PIN, memilki berbagai kelemahan seperti mudah hilang dan dicuri. Untuk mengatasi kekurangan tersebut, teknologi biometrik telah muncul sebagai alternatif salah satunya adalah *keystroke dynamics*. *Keystroke dynamics* dapat menjadi lapisan keamanan tambahan dalam melakukan autentikasi atau verifikasi menggunakan pola pengetikan pengguna. Penelitian ini berfokus pada eksplorasi metode *Distance Enhanced Flight-Time* (DEFT) untuk membangun sistem *keystroke dynamics-based authentication* (KDA). DEFT merupakan metode ekstraksi fitur *keystroke dynamic* yang menggabungkan fitur waktu penekanan tombol dengan jarak antara tombol pada *keyboard*. Sistem KDA dibangun menggunakan XGBoost sebagai model klasifikasi biner untuk melakukan autentikasi pengguna berdasarkan fitur DEFT. Berdasarkan hasil penelitian menggunakan Biomey *keystroke dataset,* sistem yang dibangun berhasil mencapai rata-rata FAR 4.71% dan FRR 14.59%.**

***Kata kunci*—biometrik, *keystroke dynamics*, autentikasi, DEFT, XGBoost**

I. PENDAHULUAN

Perkembangan teknologi yang pesat selama beberapa dekade terakhir telah berkontribusi dalam peningkatan penyimpanan dan akses data pada perangkat digital. Akses tidak sah terhadap data tersebut dapat menyebabkan kerugian finansial, dan kebocoran data sensitif yang mengancam keamanan informasi [1]. Akibatnya, kebutuhan akan sistem perlindungan data yang efisien menjadi semakin mendesak. Autentikasi merupakan aspek penting dalam menjaga keamanan data di antara berbagai metode perlindungan data. Saat ini, sistem autentikasi sering bergantung pada sesuatu yang kita ketahui dan/atau miliki seperti penggunaan kata sandi, *personal identification number* (PIN), serta kartu identitas [2, 3, 4]. Namun autentikasi jenis ini dapat dengan mudah hilang, dimanipulasi, atau dicuri sehingga menimbulkan berbagai risiko keamanan. [2, 4].

Untuk mengatasi kekurangan tersebut teknologi biometrik banyak dikembangkan sebagai alternatif maupun lapisan keamanan tambahan guna melengkapi sistem autentikasi yang sudah ada [1, 4]. Biometrik memanfaatkan karakteristik fisik dan perilaku unik manusia untuk mengidentifikasi individu yang berbeda. Karakteristik biometrik yang sulit untuk dipalsukan atau dicuri menawarkan tingkat keamanan yang lebih tinggi [5]. Biometrik fisik seperti sidik jari, pengenalan wajah, dan iris memiliki keterbatasan dalam situasi tertentu, seperti kondisi pencahayaan buruk atau cedera pada bagian tubuh yang dipindai [6]. Selain itu, sistem autentikasi berbasis biometrik fisik memerlukan perangkat keras tambahan yang tentunya menambah biaya dan kompleksitas proses autentikasi pengguna [3, 6]. Sebagai alternatif, biometrik berbasis perilaku *keystroke dynamics* yang merupakan pola pengetikan individu pada perangkat digital, menjadi solusi yang lebih praktis dan ekonomis karena tidak memerlukan perangkat tambahan [2, 3, 6].

Sejumlah penelitian terkait *keystroke dynamics* untuk mengidentifikasi dan autentikasi pengguna telah dilakukan menggunakan berbagai jenis fitur dan algoritma [7, 8, 9]. Penelitian [7] melakukan autentikasi pengguna yang menerapkan ekstraksi fitur dengan membagi *keyboard* menjadi tiga bagian, yaitu tombol yang ditekan tangan kiri (L), tombol yang ditekan tangan kanan (R), dan tombol spasi (S). Penelitian ini menunjukkan metode yang digunakan dapat meningkatkan akurasi autentikasi pada tiga perangkat berbeda, PC *keyboard*, *soft keyboard*, dan *touch keyboard*. Sementara itu, penelitian [8] membandingkan kinerja berbagai model *machine learning* untuk autentikasi pengguna berbasis *keystroke dynamics*, di mana model XGBoost menunjukkan performa terbaik dengan akurasi mencapai 96,39%. Penelitian [9] memperkenalkan metode ekstraksi fitur *distance enhanced flight-time* (DEFT) dengan memperhitungkan jarak antara tombol pada *keyboard* dalam mengukur waktu antara tombol yang ditekan. Penelitian tersebut menggabungkan fitur DEFT dengan fitur-fitur yang digunakan dalam penelitian sebelumnya, dan hasilnya menunjukkan peningkatan akurasi autentikasi yang

signifikan. Penelitian ini berhasil mencapai akurasi di atas 99% dan *equal error rate* (EER) di bawah 10% pada perangkat *desktop*, *mobile*, dan *tablet*.

Topik masalah dalam penelitian ini berfokus pada dua hal. Pertama, bagaimana cara mengimplementasikan metode DEFT pada sistem autentikasi pengguna berbasis *keystroke*. Kedua, bagaimana performa metode DEFT pada sistem KDA dengan menggunakan *dataset* Biomey. Sistem biometrik yang dikembangkan berfokus pada autentikasi atau verifikasi, sehingga hasil autentikasi akan menentukan apakah pengguna adalah pengguna sah (*genuine*) atau tidak sah (*impostor*). Penelitian ini menggunakan pendekatan yang berbeda dari penelitian sebelumnya [9], dengan mempertimbangkan tidak hanya pola pengetikan pada masing-masing tangan tetapi juga pola pengetikan saat pengguna melakukan perpindahan tangan sehingga memberikan cakupan yang lebih luas dan lebih representatif terhadap variasi pola pengetikan. *Dataset* yang digunakan dalam penelitian ini juga berbeda, di mana *dataset* yang digunakan dalam penelitian ini adalah Biomey *keystroke dataset*.

Tujuan dari penelitian ini adalah untuk mengimplementasikan metode DEFT pada sistem KDA, serta mengevaluasi performansi sistem menggunakan *dataset* Biomey. Penelitian ini diharapkan dapat memberikan wawasan mengenai kinerja metode DEFT dalam sistem autentikasi berbasis *keystroke dynamics*.

II. KAJIAN TEORI

A. Biometrik

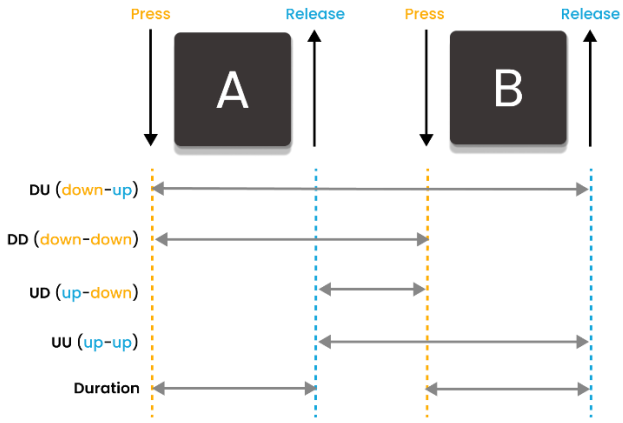
Biometrik dapat diartikan sebagai karakteristik fisik atau perilaku manusia yang dapat membedakan satu individu dengan yang lain [4]. Biometrik dibagi menjadi dua kategori utama, yaitu fisiologis dan perilaku. Biometrik fisiologis meliputi atribut fisik seperti sidik jari, wajah, dan iris mata. Sementara itu, biometrik perilaku mencakup pola-pola aktivitas atau kebiasaan seperti pola pengetikan, cara berjalan, dan tanda tangan. Sistem biometrik memiliki keunggulan dibandingkan metode keamanan tradisional seperti kata sandi dan PIN, karena tidak mudah dicuri, hilang atau dibagikan [4, 10]. Sistem biometrik juga meningkatkan kenyamanan pengguna dengan mengurangi kebutuhan untuk membuat dan mengingat kata sandi.

Sistem biometrik umumnya memiliki dua tahapan, yakni tahap pendaftaran dan tahap pengenalan [11]. Tahap pendaftaran, sistem menerima data biometrik pengguna sebagai input, kemudian melakukan *preprocessing* data dan akhirnya mengekstrak fitur dari *input* tersebut. Fitur yang diekstraksi ini akan disimpan dalam database. Tahap pengenalan, sistem menerima data biometrik baru pengguna dan kemudian melakukan ekstraksi fitur seperti pada tahap pendaftaran. Setelah ekstraksi fitur selesai, sistem akan membandingkan fitur yang diekstraksi dengan fitur yang ada di database untuk menghitung skor kemiripan dan melakukan identifikasi atau autentik

B. *Keystroke Dynamics*

*Keystroke dynamic*s merupakan proses pengukuran dan penilaian pola pengetikan individu pada *keyboard* fisik maupun *virtual* [9]. Informasi yang dihasilkan dari *keystroke dynamics* ini bersifat unik sehingga dapat digunakan untuk identifikasi dan autentikasi pengguna. Berbeda dengan sistem autentikasi berbasis biometrik fisiologis seperti iris dan sidik jari yang membutuhkan perangkat keras khusus, *keystroke dynamics* dapat sepenuhnya diimplementasikan melalui perangkat lunak [3]. Tidak hanya dapat mengurangi biaya penerapan, *keystroke dynamics* juga menawarkan cara untuk terus melakukan autentikasi identitas pengguna. Selama interaksi pengguna dengan sistem melalui perangkat *input* berlangsung, *keystroke dynamics* dapat terus dipantau dan divalidasi ulang.

Fitur *keystroke dynamics* umumnya diekstraksi menggunakan informasi waktu dari aktivitas menekan, menahan dan melepaskan tombol pada *keyboard* [12]. Digraf (n-graf dengan) yaitu *latency* antara dua penekanan tombol yang berurutan, sering digunakan sebagai fitur *keystroke dynamic*. [6, 7, 12]. Berdasarkan fitur digraf tersebut, terdapat lima jenis informasi atau fitur *keystroke dynamic* yang dapat diekstraksi [6], yaitu duration, *down-down* (DD), *up-up* (UU), *up-down* (UD), dan *down-up* (DU). *Duration* adalah waktu antara menekan dan melepaskan satu tombol. DD adalah waktu antara menekan satu tombol dan menekan tombol berikutnya, sedangkan UU adalah waktu antara melepaskan satu tombol dan melepaskan tombol berikutnya. UD adalah waktu antara melepaskan satu tombol dan menekan tombol berikutnya, sedangkan waktu DU adalah waktu antara menekan satu tombol dan melepaskan tombol berikutnya. Gambar 1 menunjukkan ilustrasi lima fitur *keystroke dynamic* menggunakan digraf.



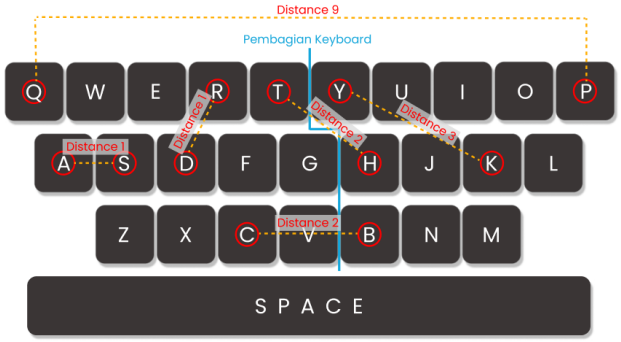
GAMBAR 1
FITUR *KEYSTROKE DYNAMIC* MENGGUNAKAN *DIGRAF*

C. *Distance Enhanced Flight-Time* (DEFT)

DEFT adalah metode ekstraksi fitur dalam *keystroke dynamics* yang diperkenalkan pada penelitian [9]. DEFT berfokus pada pengukuran jarak antara tombol-tombol pada *keyboard* dan waktu yang dibutuhkan untuk berpindah dari satu tombol ke tombol berikutnya saat mengetik. Jarak dihitung menggunakan pemisahan antar tombol yang ditentukan berdasarkan *keyboard* yang digunakan.

Gambar 2 menunjukkan jarak antara beberapa pasangan tombol. Jarak antara pasangan tombol yang dihitung berkisar dari nol hingga sembilan, dengan nol menunjukkan penekanan tombol yang sama, sedangkan sembilan menunjukkan jarak terjauh antara dua tombol, khususnya 'Q' dan 'P'. Metode DEFT juga melakukan pembagian *keyboard* seperti yang ditandai garis biru pada Gambar 2. Pembagian

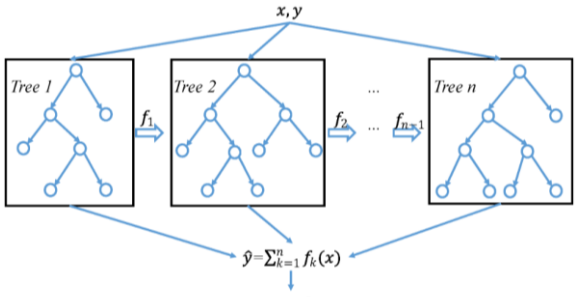
tersebut berdasarkan tombol yang diketik dengan tangan kiri dan tombol yang diketik dengan tangan kanan.



GAMBAR 2
STRUKTUR *KEYBOARD* APLIKASI BIOMEY

D. *eXtreme Gradient Boosting* (XGBoost)

XGBoost merupakan metode *gradient-boosting decision tree* yang sering digunakan untuk membuat model prediksi untuk masalah regresi dan klasifikasi [18]. XGBoost adalah metode *ensemble learning* yang bekerja dengan cara menggabungkan beberapa model yang lemah menjadi satu model yang kuat di mana setiap model baru dibuat untuk memperbaiki sisa kesalahan (*residual error*) dari model-model sebelumnya, kemudian digabungkan untuk menghasilkan prediksi akhir [18].



GAMBAR 3
ARSITEKTUR XGBOOST [14]

$$\hat{y}_i = \sum_{k=1}^{n} f_k(x_i), f_k \in F \quad (1)$$

Persamaan (1) menjelaskan bagaimana XGBoost melakukan prediksi dengan menjumlahkan *output* atau prediksi dari beberapa model (pohon). Di mana $\hat{y}_i$ merupakan *output* dari model XGBoost untuk *input* x_i , n adalah jumlah model atau pohon yang digunakan dalam *ensemble learning* $f_k(x_i)$, adalah *output* atau prediksi pohon ke-k untuk *input* x_i dalam ruang fungsi *ensemble* F [13 ,14].

E. *False Acceptance Rate* (FAR) dan *False Rejection Rate (FRR)*

False Acceptance Rate (FAR) merupakan persentase perbandingan jumlah pengguna tidak sah (*impostor*) yang dikenali sebagai penggunah sah (*genuine*) terhadap jumlah total *impostor* yang mengakses sistem [11], seperti yang ditunjukkan pada Persamaan (2), sedangkan, *False Rejection Rate* (FRR) merupakan rasio antara pengguna *genuine* yang dikenali sebagai *impostor* terhadap jumlah total pengguna sah yang mengakses sistem, seperti yang ditunjukkan pada Persamaan (3).

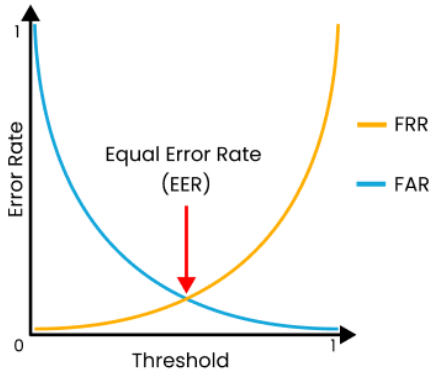
$$FAR = \frac{Number\ of\ false\ acceptances}{Total\ number\ of\ impostor\ attempts} \quad (2)$$

$$FRR = \frac{Number\ of\ false\ rejections}{Total\ number\ of\ genuine\ attempts} \quad (3)$$

Nilai ideal untuk FAR dan FRR adalah 0%, yang artinya memberikan akses kepada semua *genuine* sementara semua *impostor* ditolak [15]. Sebuah sistem dengan 0% FAR dianggap memiliki tingkat keamanan tertinggi, sementara sistem dengan 0% FRR sangat mudah digunakan. Keseimbangan antara keamanan dan kemudahan penggunaan sangat penting, karena sistem dengan keamanan tinggi mungkin menolak terlalu banyak pengguna sah, sehingga tidak mudah digunakan. Sebaliknya, sistem yang terlalu mudah digunakan menjadi kurang aman.

F. *Equal Error Rate* (EER)

Nilai FAR dan FRR dapat digambarkan ke bentuk grafik dengan kurva nilai *error* pada berbagai nilai *threshold*. Pada grafik tersebut, titik perpotongan kedua kurva disebut EER [16]. EER merupakan titik di mana FAR dan FRR bernilai sama atau memiliki selisih terkecil pada rentang nilai *threshold* yang digunakan. Sama halnya dengan FAR dan FRR, semakin rendah nilai EER, semakin baik pula performansi model yang digunakan. Kurva relasi antara FAR, FRR dan EER dapat dilihat pada Gambar 4.



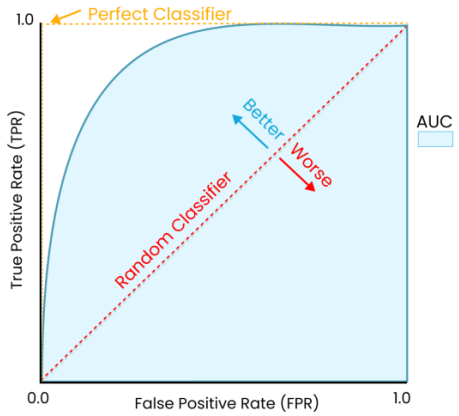
GAMBAR 4
GRAFIK RELASI ANTARA FAR, FRR, DAN EER

G. ROC-AUC

Receiver operating characteristic (ROC) menggambarkan kurva *true positive rate* (TPR) terhadap *false positive rate* (FPR) atau FAR pada berbagai nilai *threshold*. *Area under the curve* (AUC) mengukur seberapa baik model dapat membedakan antara kelas pada model klasifikasi. ROC-AUC digunakan untuk menilai kinerja model klasifikasi pada berbagai *threshold* [15]. Nilai AUC yang lebih tinggi menunjukkan model yang lebih baik. Model yang baik akan memiliki titik semakin dekat dengan sudut kiri atas atau koordinat (0.0, 1.0) yang menunjukkan tidak ada *false accept* atau *false reject* [15].

$$TPR = \frac{Number\ of\ true\ acceptances}{Total\ number\ of\ genuine\ attempts} \quad (4)$$

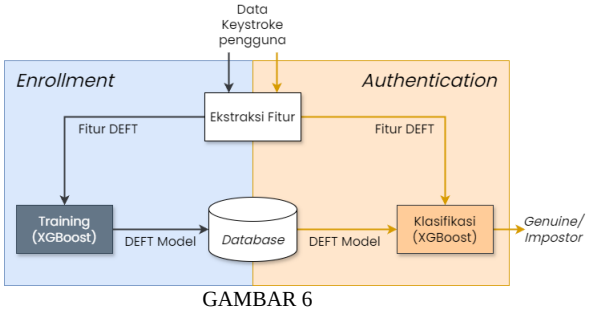
$$FPR = \frac{Number\ of\ false\ acceptances}{Total\ number\ of\ impostor\ attempts} \quad (5)$$



GAMBAR 5
GRAFIK ROC-AUC

III. METODE

Sistem KDA yang dibangun dalam penelitian ini melibatkan dua tahapan utama, *enrollment* dan *authentication*. Pada tahap *enrollment*, *input* yang digunakan adalah data *keystroke* dari pengguna, yang mencakup waktu penekanan dan pelepasan tombol saat mengetik. Data ini diolah melalui ekstraksi fitur menggunakan metode DEFT, menghasilkan fitur-fitur *keystroke* yang mencerminkan karakteristik unik pengguna saat mengetik. Fitur-fitur ini kemudian diolah untuk membangun *template keystroke* yang unik untuk setiap pengguna. Ketika tahap *authentication*, *input* yang digunakan adalah data *keystroke* baru yang diambil saat pengguna mencoba mengakses sistem. Data ini diproses melalui ekstraksi fitur yang sama seperti pada tahap *enrollment*, menghasilkan fitur *keystroke* baru. Fitur-fitur ini kemudian dibandingkan dengan *template* yang ada di *database*, menggunakan model klasifikasi biner yang telah dilatih sebelumnya. *Output* dari proses ini adalah nilai probabilitas yang menunjukkan seberapa besar kemungkinan *keystroke* tersebut berasal dari pengguna sah. Jika probabilitas tersebut melebihi nilai *threshold* yang disimpan, sistem akan mengizinkan akses pengguna sebagai *genuine*. Sebaliknya, jika nilai probabilitas lebih rendah dari *threshold*, sistem akan menolak akses dan mengenali pengguna sebagai *impostor*. Gambaran umum rancangan sistem KDA yang dibangun ditunjukkan pada Gambar 6.



GAMBAR 6
GAMBARAN UMUM RANCANGAN SISTEM AUTENTIKASI BERBASIS *KEYSTROKE DYNAMIC*

A. *Dataset*

Penelitian ini menggunakan Biomey *keystroke dataset* [19], yang merupakan hasil pengumpulan data *keystroke* menggunakan aplikasi khusus berbasis android pada perangkat *smartphone*. Pengumpulan data melibatkan 40 partisipan dalam 30 sesi pengambilan dengan total data sekitar 500.000 *keystroke*. Contoh data dari *dataset* yang digunakan dapat dilihat pada Tabel 1.

TABEL 1
CONTOH *DATASET*

UID	SID	Kalimat	*Press (ms)*	*Release (ms)*	*Key*
10001	1	1	67642	67723	h
10001	1	1	67905	67990	e
10001	1	1	68036	68124	l
10001	1	1	68706	68775	l
10001	1	1	68891	68958	o
10001	1	1	69039	69125	SPACE
10001	1	1	69190	69274	w
10001	1	1	69372	69458	o
10001	1	1	69906	69960	r
10001	1	1	70023	70127	l
10001	1	1	70158	70261	d

B. *Enrollment*

Proses enrollment mencakup proses pembentukan digraf dan ekstraksi fitur menggunakan metode DEFT, di mana data *keystroke* diubah menjadi fitur yang dapat digunakan untuk membedakan antara pengguna *genuine* dan *impostor*. Hasil dari ekstraksi fitur tersebut kemudian digunakan untuk membentuk *template keystroke*. *Template* ini merupakan representasi pola pengetikan unik dari setiap pengguna, yang nantinya akan disimpan dalam *database*. *Template* yang tersimpan ini akan digunakan sebagai acuan selama proses autentikasi untuk menentukan identitas pengguna.

1. Pembentukan Digraf

Pembentukan digraf dilakukan dengan membuat pasangan dua tombol berurutan yang ditekan oleh pengguna. Informasi waktu terkait dengan penekanan dan pelepasan tombol-tombol ini dapat digunakan untuk mendapatkan fitur *latency* digraf, yaitu DD, UU, DU, UD, dan *duration*. Untuk setiap pasangan digraf (k_i, k_{i+1}) fitur *latency* waktu dapat dihitung menggunakan persamaan sebagai berikut.

$$DD_{k_i,k_{i+1}} = press\_time_{k_{i+1}} - press\_time_{k_i} \quad (6)$$

$$UU_{k_i,k_{i+1}} = release\_time_{k_{i+1}} - release\_time_{k_i} \quad (7)$$

$$DU_{k_i,k_{i+1}} = press\_time_{k_{i+1}} - release\_time_{k_i} \quad (8)$$

$$UD_{k_i,k_{i+1}} = release\_time_{k_{i+1}} - press\_time_{k_i} \quad (9)$$

$$Duration_{k_i,k_{i+1}} = DU_{k_i,k_{i+1}} - UD_{k_i,k_{i+1}} \quad (10)$$

2. Ekstraksi Fitur DEFT

Ekstraksi fitur menggunakan metode DEFT diawali dengan mengelompokkan data berdasarkan posisi tombol pada pasangan digraf, kemudian dilanjutkan dengan mengukur jarak antar tombol pada *keyboard*. Pembagian kelompok dilakukan berdasarkan tombol yang diketik menggunakan tangan kiri (L) dan tombol yang diketik menggunakan tangan kanan (R). Kelompok yang digunakan hanya mencakup LL dan RR pada penelitian [9]. Untuk penelitian ini, kelompok yang digunakan diperluas menjadi LL, RR, LR, dan RL, sedangkan kelompok data spasi dihiraukan. Tabel 2 menunjukkan contoh hasil pengelompokan dan pengukuran jarak antar pasangan digraf

berdasarkan struktur *keyboard* pada aplikasi Biomey [19] yang ditunjukkan pada Gambar 2.

TABEL 2
HASIL PEMBENTUKAN FITUR DIGRAF DAN DEFT

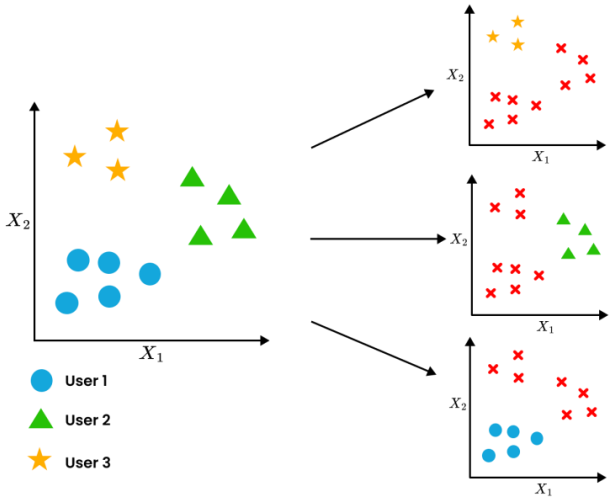
Digraf							DEFT	
DD	***UU***	***DU***	***UD***	***Duration***	***Key 1***	***Key 2***	***Group***	***Distance***
263	267	348	182	166	h	e	RL	4
131	134	219	46	173	e	l	LR	7
670	651	739	582	157	l	l	RR	0
185	183	252	116	136	l	o	RR	1
182	184	268	98	170	w	o	LR	7
534	502	588	448	140	o	r	RL	5
117	167	221	63	158	r	l	LR	6
136	134	238	31	207	l	d	RL	6

3. *Training*

Untuk tahap *training*, data hasil ekstraksi fitur digunakan untuk membangun dan melatih model klasifikasi XGBoost. D*ataset* yang digunakan terdiri dari 40 pengguna atau kelas. Karena sistem yang dibangun adalah sistem autentikasi, yang bertujuan untuk membedakan antara pengguna yang sah (*genuine*) dan tidak sah (*impostor*), dilakukan penerapan strategi *one-vs-rest* (OVR) untuk mengubah masalah klasifikasi *multiclass* menjadi serangkaian model klasifikasi biner.

Strategi OVR ini memungkinkan setiap pengguna diperlakukan sebagai kelas *genuine* tersendiri, sementara semua pengguna lainnya digabungkan menjadi satu kelas *impostor*, dengan kata lain, setiap pengguna memiliki model klasifikasi biner tersendiri, yang dirancang untuk membedakan *keystroke* pengguna tersebut dari *keystroke* pengguna lainnya. Misalnya, model klasifikasi biner untuk pengguna A dilatih untuk mengidentifikasi apakah data *keystroke* yang diberikan berasal dari pengguna A atau dari pengguna lain di *dataset*. Proses ini diulangi untuk semua pengguna, sehingga dihasilkan total 40 model klasifikasi biner sesuai dengan jumlah pengguna atau kelas dalam *dataset*.

Untuk tahap *training* ini digunakan berbagai rentang nilai *threshold* dalam melakukan klasifikasi untuk mendapatkan *threshold* optimal bagi setiap pengguna. Hasil dari ekstraksi fitur beserta nilai *threshold* yang diperoleh kemudian disimpan ke dalam *database* sebagai *template* unik untuk masing-masing pengguna.



GAMBAR 7
STRATEGI *ONE VS REST* (OVR)

C. *Authentication*

Tahap autentikasi adalah proses pengambilan keputusan untuk mengenali apakah pengguna sah (*genuine*) atau tidak sah (*impostor*). Autentikasi dimulai dengan melakukan ekstraksi fitur pada data *keystroke* baru pengguna, kemudian data hasil ekstraksi fitur di klasifikasi dengan model dari proses *enrollment* yang menghasilkan nilai probabilitas kemungkinan pengguna merupakan kelas *genuine*. Probabilitas tersebut kemudian akan dibandingkan dengan nilai *threshold* yang ada pada *database*. Jika nilai probabilitas lebih tinggi atau sama dengan nilai *threshold*, maka model klasifikasi akan mengenali pengguna sebagai *genuine*. Sebaliknya jika nilai probabilitas lebih rendah dari nilai *threshold*, maka pengguna akan dikenali sebagai *impostor*. Tabel 3 menunjukkan contoh hasil autentikasi dari sistem yang dibangun.

TABEL 3
CONTOH HASIL AUTENTIKASI

Label	**Probabilitas**	***Threshold***	**Hasil Autentikasi**
Genuine	0.8	0.5	*Genuine*
Genuine	0.49	0.5	*Impostor*
Genuine	0.1	0.5	*Impostor*
Impostor	0.05	0.5	*Impostor*
Impostor	0.1	0.5	*Impostor*
Impostor	0.6	0.5	*Genuine*

IV. HASIL DAN PEMBAHASAN

Evaluasi performansi sistem diukur dengan menggunakan dua metrik performansi yaitu FAR, FRR dan EER serta ROC-AUC. EER didapatkan dengan mencari selisih terkecil antara FAR dan FRR pada berbagai threshold yang dapat memberikan indikasi seberapa baik sistem dapat menyeimbangkan *error* dalam autentikasi pengguna, dengan nilai EER yang lebih rendah menunjukkan kinerja yang lebih baik. Selain itu, ROC-AUC digunakan untuk menilai performa model secara keseluruhan yang menggambarkan *trade-off* antara TPR dan FPR pada berbagai nilai *threshold*. Nilai ROC-AUC yang lebih tinggi menunjukkan bahwa model memiliki kemampuan yang lebih baik dalam membedakan antara kelas *genuine* dan *impostor*. *Dataset* dibagi dengan rasio 80% untuk data *training* dan 20% untuk data *testing*. Proses evaluasi dilakukan melalui tiga skenario pengujian sebagai berikut.

A. Skenario Pengujian Pertama

TABEL 4
AUC DAN EER SKENARIO PERTAMA MASING-MASING FITUR WAKTU DIGRAF (LL-RR)

Fold	**DU**		**UU**		**UD**		**DD**		***Duration***	
	AUC	**EER(%)**	**AUC**	**EER(%)**	**AUC**	**EER(%)**	**AUC**	**EER(%)**	**AUC**	**EER(%)**
1	0.778	24.5	0.739	28	0.728	29.9	0.743	28.3	0.827	20.5
2	0.767	26.6	0.751	27.5	0.721	30.4	0.774	24.5	0.807	21.7
3	0.776	25.2	0.761	26.5	0.751	26.9	0.779	24.9	0.84	18.5
4	0.775	25	0.754	27.5	0.737	28.1	0.749	27.7	0.818	21.2
5	0.76	26.4	0.75	26.9	0.762	25.9	0.754	26.3	0.843	18.1
Rata-rata	0.771	25.5	0.751	27.3	0.74	28.2	0.76	26.3	0.827	20

TABEL 5
AUC DAN EER SKENARIO PERTAMA MASING-MASING FITUR WAKTU DIGRAF (LL-LR-RL-RR)

Fold	**DU**		**UU**		**UD**		**DD**		***Duration***	
	AUC	**EER(%)**	**AUC**	**EER(%)**	**AUC**	**EER(%)**	**AUC**	**EER(%)**	**AUC**	**EER(%)**
1	0.847	18	0.854	17.5	0.84	19.3	0.838	18.4	0.861	16.3
2	0.837	18.4	0.828	19.6	0.815	20.8	0.837	18.5	0.863	16.4
3	0.849	17.9	0.871	16	0.877	14.6	0.886	14.1	0.901	12.4
4	0.89	13.1	0.859	16.9	0.84	18.6	0.842	18.5	0.893	14
5	0.836	19	0.827	19.9	0.823	20	0.833	19.1	0.874	14.6
Rata-rata	0.852	17.3	0.848	18	0.839	18.7	0.847	17.7	0.878	14.7

TABEL 6
AUC DAN EER SKENARIO PERTAMA DENGAN PENGGABUNGAN FITUR WAKTU DIGRAF

Fold	**LL-RR**		**LL-LR-RL-RR**	
	AUC	**EER(%)**	**AUC**	**EER(%)**
1	0.895	12.86	0.919	10.28
2	0.888	14.01	0.915	10.53
3	0.892	13.65	0.942	7.7
4	0.905	12.2	0.929	8.82
5	0.907	11.93	0.911	11.08
Rata-rata	0.897	12.9	0.923	9.7

Dalam skenario pengujian pertama, dilakukan perbandingan performansi masing-masing fitur waktu digraf (UD, DD, DU, UU, *duration)* serta penggabungan fitur waktu tersebut dengan dua kombinasi kelompok yang berbeda, yaitu (LL-RR) dan (LL-LR-RL-RR). Pengujian dilakukan menggunakan data *training* dengan menerapkan *5-fold cross-validation* di mana nilai EER dan AUC untuk setiap *fold* merupakan rata-rata nilai dari 40 model klasifikasi biner. Selain itu, s*tratified k-fold* digunakan untuk memastikan bahwa setiap *fold* dalam proses *cross-validation* memiliki distribusi data pengguna yang seimbang.

B. Skenario Pengujian Kedua

Pada skenario pengujian kedua, dilakukan *tuning hyperparameter* pada dua model XGBoost yang dibangun menggunakan kombinasi kelompok yang berbeda, di mana fitur yang digunakan adalah penggabungan fitur waktu digraf. Proses *tuning hyperparameter* dilakukan dengan cara yang sama dengan skenario pengujian pertama. Parameter yang diuji adalah *n_estimators*, *learning_rate*, *max_depth*, *min_child_weight*, *scale_pos_weight* dan *gamma*.

C. Skenario Pengujian Ketiga

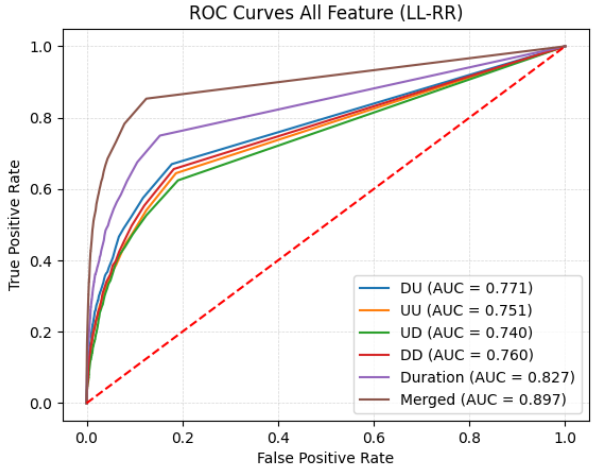
Pada skenario ketiga, fokus pengujian adalah mengevaluasi performa model serta *threshold* masing-masing pengguna yang didapatkan dari proses *enrollment*. Setiap fitur waktu digraf (UD, DD, DU, UU, dan *duration*) diuji menggunakan model XGBoost dengan parameter *default*. Pengujian ini bertujuan untuk menilai performa dasar dari masing-masing fitur waktu dalam mendeteksi pengguna *genuine* dan *impostor*, sedangkan untuk penggabungan kelima fitur waktu, dilakukan perbandingan antara model XGBoost dengan parameter *default* dan model yang telah

melalui proses *tuning hyperparameter* pada skenario pengujian kedua. Tujuan dari perbandingan ini adalah untuk mengevaluasi sejauh mana proses *tuning hyperparameter* dapat meningkatkan performa sistem KDA dibandingkan dengan penggunaan parameter *default*. Evaluasi performansi dilakukan dengan mengukur FAR dan FRR dalam proses autentikasi pada data testing untuk setiap pengguna. Selain itu, pengukuran waktu juga dilakukan untuk melihat kinerja sistem dalam melakukan autentikasi terhadap 40 pengguna berbeda, di mana masing-masing pengguna diuji menggunakan data *testing* yang telah disiapkan.

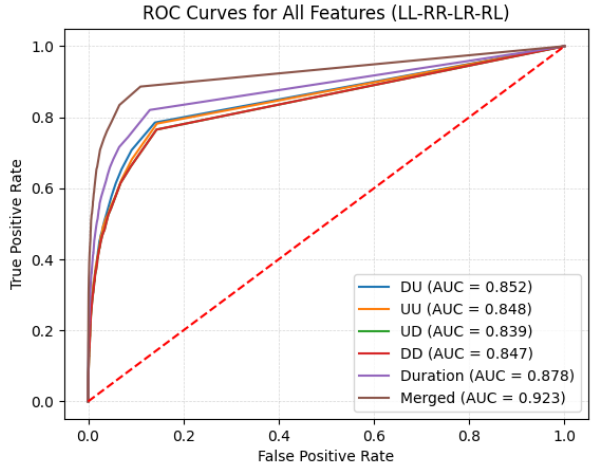
D. Hasil Pengujian

1. Skenario Pengujian Pertama

Pengujian perbandingan masing-masing fitur dilakukan dengan mengukur performansi model pada setiap *fold* menggunakan metrik EER dan AUC. Tabel 4 dan Tabel 5 menunjukkan hasil performansi model menggunakan masing-masing fitur waktu digraf sedangkan hasil pengujian menggunakan penggabungan fitur waktu digraf ditunjukkan pada Tabel 6.



GAMBAR 8
PERBANDINGAN ROC-AUC SKENARIO PENGUJIAN PERTAMA (LL-RR)



GAMBAR 9
PERBANDINGAN ROC-AUC SKENARIO PENGUJIAN PERTAMA (LL-LR-RL-RR)

2. Skenario Pengujian Kedua

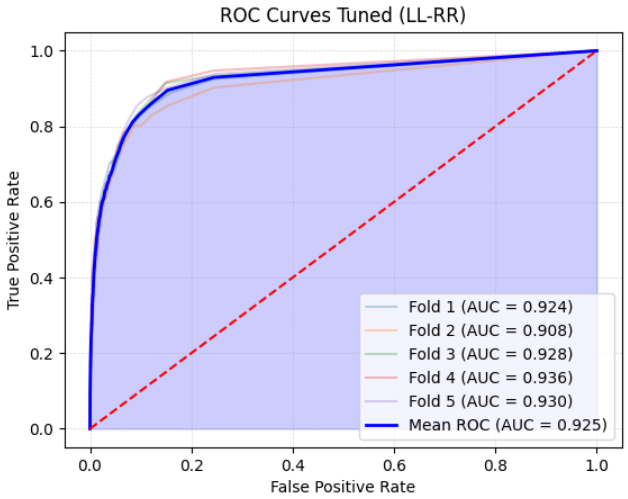
Tuning hyperparameter dilakukan dengan menguji model dengan berbagai kombinasi nilai parameter dan menggunakan nilai rata-rata EER dari *5-fold cross validation* sebagai metrik acuan. Rentang nilai yang di uji untuk setiap parameter beserta model terbaik hasil *tuning hyperparameter* untuk kedua kombinasi kelompok ditunjukkan pada Tabel 7. Tabel 8 menunjukkan hasil performansi model hasil *tuning hyperparameter* menggunakan penggabungan fitur waktu.

TABEL 7
PARAMETER MODEL HASIL *TUNING HYPERPARAMETER*

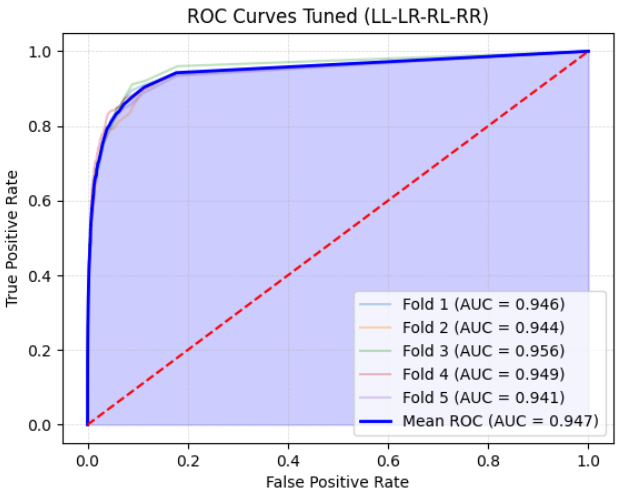
Parameter	**Nilai Yang Di Uji**	**Model Terbaik**	
		LL-RR	**LL-LR-RL-RR**
n_estimators	{50, 100, 200}	200	200
learning_rate	{0.05, 0.1, 0.2, 0.3}	0.05	0.1
scale_pos_weight	$\{0, \sqrt{39}, 39\}$	$\sqrt{39}$	39
max_depth	{3, 4, 5, 6, 7}	5	4
min_child_weight	{1, 2, 3}	2	2
gamma	{0. 0.1, 0.2}	0.2	0.1

TABEL 8
AUC DAN EER SKENARIO PENGUJIAN KEDUA MENGGUNAKAN PENGGABUNGAN FITUR WAKTU DIGRAPH

Fold	**LL-RR**		**LL-LR-RL-RR**	
	AUC	**EER(%)**	**AUC**	**EER(%)**
1	0.924	12.02	0.946	9.41
2	0.908	14.5	0.944	9.24
3	0.928	11.65	0.956	7.56
4	0.936	9.76	0.949	7.25
5	0.93	11.48	0.941	8.96
Rata-rata	0.925	11.9	0.947	8.5



GAMBAR 10
ROC-AUC MODEL HASIL SKENARIO PENGUJIAN KEDUA (LL-RR)



GAMBAR 11
ROC-AUC MODEL HASIL SKENARIO PENGUJIAN KEDUA (LL-LR-RL-RR)

3. Skenario Pengujian Ketiga

Pengujian performansi model untuk masing-masing fitur waktu individu menggunakan *threshold* dari proses *enrollment* ditunjukkan pada Tabel 9. Sedangkan hasil performansi untuk model sebelum dan sesudah *tuning hyperparameter* dengan penggabungan fitur waktu ditunjukkan pada Tabel 10.

TABEL 9
FAR DAN FRR SKENARIO PENGUJIAN KETIGA MASING-MASING FITUR WAKTU DIGRAF

Group	**Fitur**	**FAR (%)**	**FRR (%)**	***Time(s)***
LL-RR	DU	14.66	43.33	3.06
	UU	15.0	40.83	3.40
	UD	16.52	40.0	3.14
	DD	15.69	37.92	3.09
	Duration	11.41	27.08	2.76
LL-LR-RL-RR	DU	9.70	30.42	8.79
	UU	9.48	27.92	5.06
	UD	10.19	30.0	4.98
	DD	9.42	29.1	8.62
	Duration	7.98	25.42	7.67

TABEL 10
FAR DAN FRR SKENARIO PENGUJIAN KETIGA UNTUK PENGGABUNGAN FITUR WAKTU DIGRAF

Model	**LL, RR**			**LL, LR, RL, RR**		
	FAR(%)	**FRR(%)**	***Time(s)***	**FAR(%)**	**FRR(%)**	***Time(s)***
Default	7.02	26.25	10.89	4.78	15.42	21.78
Tuned	8.22	20.42	24.38	4.71	14.59	54.99

E. Analisis Hasil Pengujian

Untuk skenario pengujian pertama, hasil pengujian menunjukkan bahwa performansi kelompok (LL-LR-RL-RR) selalu lebih baik dari kelompok (LL-RR) untuk setiap fitur waktu digraf yang digunakan. Fitur *duration* menjadi fitur waktu individu dengan performansi terbaik untuk kedua kombinasi kelompok. Pengujian ini juga memperlihatkan penggabungan fitur waktu digraf memiliki hasil yang lebih baik dibandingkan dengan penggunaan fitur waktu digraf secara individu.

Skenario pengujian kedua memperlihatkan hasil proses *tuning hyperparameter* pada model XGBoost dapat meningkatkan kinerja sistem KDA. Karena *dataset* yang digunakan *imbalanced*, di mana lebih banyak data kelas negatif dibandingkan dengan kelas positif, parameter *scale_pos_weight* digunakan untuk mengubah bobot data kelas positif. *Imbalanced dataset* ini sendiri terjadi karena untuk melakukan autentikasi *dataset* yang sebelumnya berbentuk *multi-class* diubah menjadi biner dengan strategi *one vs rest*. Meskipun jumlah kombinasi parameter yang di uji relatif terbatas, proses ini mampu meningkatkan rata-rata EER dari 12.9% menjadi 11.9% untuk kelompok (LL-RR), sedangkan untuk kelompok (LL-LR-RL-RR) yang sebelumnya mendapatkan EER 9.7% berhasil mencapai EER 8.5%.

Baik pada skenario pengujian pertama maupun kedua, terlihat adanya ketidakstabilan pada nilai EER dan AUC yang dihasilkan. Salah satu faktor yang dapat berkontribusi terhadap ketidakstabilan ini adalah tidak dilakukannya *preprocessing* data dalam penelitian ini. Tanpa adanya *preprocessing* menyebabkan data anomali atau *outlier* yang mungkin ada dalam dataset tetap digunakan dalam proses pelatihan model, sehingga menyebabkan fluktuasi performa yang signifikan antar *fold*. Selain itu, *dataset* yang relatif kecil dan tidak seimbang antara kelas *genuine* dan *impostor* juga turut memperburuk ketidakstabilan tersebut.

Berdasarkan hasil skenario pengujian ketiga, pengujian pada data testing menggunakan model yang sudah dibangun berdasarkan skenario pertama dan kedua serta menerapkan *threshold* dari proses *enrollment*, menunjukkan bahwa kelompok (LL-LR-RL-RR) menghasilkan kinerja yang lebih baik namun membutuhkan waktu eksekusi yang lebih lama

dibandingkan dengan kelompok (LL-RR). Hal ini dipengaruhi oleh jumlah dimensi fitur yang digunakan, di mana kelompok (LL-RR) memiliki 55 fitur sedangkan kelompok (LL-LR-RL-dRR) mempunyai 145 fitur

V. KESIMPULAN

Pada penelitian ini, autentikasi pengguna berbasis *keystroke dynamics* dibangun menggunakan metode ekstraksi fitur DEFT dan model klasifikasi XGBoost. Hasil penelitian menunjukkan bahwa pemilihan fitur dan kelompok yang digunakan berpengaruh terhadap kinerja sistem autentikasi yang dibangun, sehingga perlu dilakukan pemilihan yang tepat. Berdasarkan hasil pengujian menggunakan nilai *threshold* terbaik dengan penggabungan fitur waktu dan kombinasi kelompok fitur (LL-RR) memperoleh FAR 8.22% dan FRR 20.42%. Sedangkan kombinasi kelompok fitur (LL-LR-RL-RR) berhasil mencapai FAR 4.71% dan FRR 14.59%. Secara keseluruhan, hasil penelitian menunjukkan bahwa penerapan metode DEFT untuk sistem autentikasi berbasis *keystroke dynamic*s mampu memberikan performa yang menjanjikan. Untuk penelitian selanjutnya, dapat dilakukan eksplorasi lebih lanjut terkait metode DEFT seperti penggunaan *dataset*, perangkat, kelompok maupun model klasifikasi yang berbeda. Selain itu, disarankan untuk melakukan *preprocessing* dataset untuk menangani data-data *outliers* yang dapat memengaruhi performa model.

REFERENSI

[1] A. Arsh, N. Kar, S. Das, and S. Deb, "Multiple Approaches Towards Authentication Using Keystroke Dynamics," *Procedia Computer Science*, vol. 235, pp. 2609–2618, 2024, doi: 10.1016/j.procs.2024.04.246.

[2] Y. Shi, X. Wang, K. Zheng, and S. Cao, "User authentication method based on keystroke dynamics and mouse dynamics using HDA," *Multimedia Systems*, vol. 29, no. 2, pp. 653–668, Apr. 2023, doi: 10.1007/s00530-022-00997-5.

[3] P. S. Teh, A. B. J. Teoh, and S. Yue, "A Survey of Keystroke Dynamics Biometrics," *The Scientific World Journal*, vol. 2013, p. e408280, Nov. 2013, doi: 10.1155/2013/408280.

[4] A. K. Jain and A. Ross, "Introduction to Biometrics," in Handbook of Biometrics, A. K. Jain, P. Flynn, and A. A. Ross, Eds., Boston, MA: *Springer US*, 2008, pp. 1–22. doi: 10.1007/978-0-387-71041-9_1.

[5] M. L. Ali, J. V. Monaco, C. C. Tappert, and M. Qiu, "Keystroke Biometric Systems for User Authentication," *Journal of Signal Processing System*, vol. 86, no. 2–3, pp. 175–190, Mar. 2017, doi: 10.1007/s11265-016-1114-9.

[6] J. Kim and P. Kang, "Freely typed keystroke dynamics-based user authentication for mobile devices based on heterogeneous features," *Pattern Recognition*, vol. 108, p. 107556, Dec. 2020, doi: 10.1016/j.patcog.2020.107556.

[7] P. Kang and S. Cho, "Keystroke dynamics-based user authentication using long and free text strings from various input devices," *Information Sciences*, vol. 308, pp. 72–93, Jul. 2015, doi: 10.1016/j.ins.2014.08.070.

[8] H.-C. Chang, J. Li, C.-S. Wu, and M. Stamp, "Machine learning and deep learning for fixed-text keystroke dynamics," *Advances in Information Security*, pp. 309–329, 2022. doi:10.1007/978-3-030-97087-1_13

[9] N. Kaluarachchi, S. Kandanaarachchi, K. Moore, and A. Arakala, "DEFT: A New Distance-Based Feature Set for Keystroke Dynamics," 2023 *International Conference of the Biometrics Special Interest Group (BIOSIG)*, pp. 1–6, Sep. 2023, doi: 10.1109/BIOSIG58226.2023.10345982.

[10] I. Buciu and A. Gacsadi, "Biometrics systems and technologies: A survey," *International Journal of Computers Communications & Control*, vol. 11, no. 3, p. 315, Mar. 2016. doi:10.15837/ijccc.2016.3.2556

[11] Z. Qin, P. Zhao, T. Zhuang, F. Deng, Y. Ding, and T. Chen, "A survey of identity recognition via data fusion and feature learning - ScienceDirect," *Information Fusion*, vol. 91, pp. 694–712, Mar. 2023, doi: 10.1016/j.inffus.2022.10.032.

[12] Y. Zhong and Y. Deng, "A Survey on Keystroke Dynamics Biometrics: Approaches, Advances, and Evaluations," in Gate to Computer Science and Research, 1st ed., vol. 2, Y. Zhong and Y. Deng, Eds., Science Gate Publishing P.C., 2015, pp. 1–22. doi: 10.15579/gcsr.vol2.ch1.

[13] G. J. Krishna, H. Jaiswal, P. S. R. Teja, and V. Ravi, "Keystroke based User Identification with XGBoost," *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*, pp. 1369–1374, Oct. 2019, doi: 10.1109/TENCON.2019.8929453.

[14] Y. Wang, Z. Pan, J. Zheng, L. Qian, and M. Li, "A Hybrid Ensemble method for Pulsar Candidate Classification," Astrophys Space Sci, vol. 364, no. 8, p. 139, Aug. 2019, doi: 10.1007/s10509-019-3602-4.

[15] R. Shadman, A. A. Wahab, M. Manno, M. Lukaszewski, D. Hou, and F. Hussain, "Keystroke Dynamics: Concepts, Techniques, and Applications," Mar. 08, 2023, arXiv: arXiv:2303.04605. doi: 10.48550/arXiv.2303.04605.

[16] A. Rahman et al., "Multimodal EEG and Keystroke Dynamics Based Biometric System Using Machine Learning Algorithms," *IEEE Access*, vol. 9, pp. 94625–94643, 2021, doi: 10.1109/ACCESS.2021.3092840.

[17] A.-C. Iapa and V.-I. Cretu, "Modified Distance Metric That Generates Better Performance For The Authentication Algorithm Based On Free-Text Keystroke Dynamics," in *2021 IEEE 15th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, May 2021, pp. 000455–000460. doi: 10.1109/SACI51354.2021.9465601.

[18] A. Asselman, M. Khaldi, and S. Aammou, "Enhancing the prediction of student performance based on the

machine learning XGBoost algorithm," *Interactive Learning Environments*, vol. 31, no. 6, pp. 3360–3379, Aug. 2023, doi: 10.1080/10494820.2021.1928235.

[19] I. W. A. Wahyudi, Autentikasi Pengguna Berbasiskan Biometrik Keystroke Menggunakan Instance-Based Tail Area Density. Universitas Telkom, S1 Informatika, 2023. Accessed: Apr. 24, 2024. [Online]. Available: https://openlibrary.telkomuniversity.ac.id/pustaka/197125/autentikasi-pengguna-berbasiskan-biometrik-keystroke-menggunakan-instance-based-tail-area-density.html