

BAB 1 PENDAHULUAN

1.1. Latar Belakang

Dewasa ini, keamanan pengguna menjadi sangat penting, terutama dalam hal mengakses informasi dan layanan daring [1]. Identifikasi berbasis biometrik dianggap sebagai solusi yang lebih aman dan mudah dibandingkan dengan metode tradisional seperti kata sandi atau (*personal identification number*) PIN [2]. Dengan meningkatnya ancaman keamanan seperti akses ilegal, dibutuhkan pendekatan yang lebih baik dalam pengamanan pengguna [1], [3], [4]. Biometrik adalah konsep yang merujuk pada beragam karakteristik fisik dan perilaku manusia yang dapat digunakan untuk mengidentifikasi individu tetapi tidak terbatas pada sidik jari, ciri wajah, iris, dan suara [5], [6]. Dalam konteks ini, identifikasi berbasis *keystroke* hadir sebagai alternatif yang menjanjikan, mengingat keunikan pola mengetik individu yang dapat dijadikan sebagai identitas unik. Dibandingkan dengan metode tradisional seperti kata sandi, pendekatan biometrik *keystroke* memiliki tingkat keamanan yang lebih tinggi karena sulit untuk diretas atau disalin oleh pihak yang tidak sah [2]. Oleh karena itu, memahami lebih dalam bagaimana *keystroke* bisa digunakan sebagai identitas pengguna menjadi hal yang penting, terutama untuk menghadapi ancaman keamanan yang semakin beragam di era digital ini.

Sejumlah penelitian telah dilakukan untuk mengeksplorasi *keystroke dynamics-based identification*. Contohnya adalah metode yang digunakan oleh Ioannis Tsimperidis [7], yang memanfaatkan *dataset* IKDD dalam pengujian klasifikasi pengguna berbasis *keystroke*. Penelitian ini menggunakan algoritma *Support Vector Machine* (SVM) untuk membangun model klasifikasi. Hasil evaluasi menunjukkan bahwa model SVM mampu mencapai *F1-score* terbaik sebesar 0,76.

Penelitian lain dilakukan oleh B. Radha Krishna [8], yang mengkaji penggunaan algoritma *K-Nearest Neighbors* (KNN) untuk tugas klasifikasi berbasis *keystroke dynamics*. Dalam penelitian ini, pendekatan KNN diterapkan pada data *keystroke*, dan hasil evaluasi menggunakan metrik *F1-score* menunjukkan bahwa algoritma ini mencapai *F1-score* terbaik sebesar 0,62. Hal ini memperlihatkan bahwa KNN juga dapat digunakan untuk tugas klasifikasi. Meskipun penelitian menggunakan KNN memiliki hasil yang cukup baik, namun banyak kekurangan yang dimiliki oleh algoritma KNN [9]. Pada algoritma KNN nilai ukuran tetangga (k) memiliki sensitivitas yang sangat berpengaruh [9], [10]. Dimana nilai k yang tidak tepat dapat memengaruhi kinerja algoritma, terutama dalam menghadapi *outlier*, apalagi ketika KNN dihadapi dengan volume *dataset* yang besar [9], [10]. Pada KNN tidak selalu menghasilkan klasifikasi yang *robust* terhadap *outlier*, karena aturan mayoritas suara (nilai) dari tetangga terdekat dapat dipengaruhi oleh data yang tidak representatif [9]. Untuk mengatasi masalah-masalah ini, penggunaan *multi-voter multi-commission nearest neighbor classifier* (MVMCNN), yang merupakan modifikasi dari KNN, dapat menjadi solusi yang potensial. Dari total 30 *dataset* yang diuji, 17 di antaranya menunjukkan hasil yang bagus ketika menggunakan MVMCNN [9]. Hal ini menunjukkan sekitar 57% dari *dataset* yang dievaluasi berhasil memberikan hasil yang baik dengan model MVMCNN daripada KNN, *Local Mean-based Probabilistic Neural Network* (LMPNN), dan *Bonferroni Mean Fuzzy K-Nearest Neighbors* (BM-FKNN) [9].

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas maka dirumuskan masalah yaitu bagaimana mengimplementasikan MVMCNN pada sistem identifikasi pengguna dengan *keystroke* biometrik untuk meningkatkan keamanan pengguna. Permasalahan selanjutnya adalah bagaimana performansi

klasifikasi menggunakan algoritma MVMCNN yang diterapkan pada Biomey *Keystroke Dataset*.

1.3. Tujuan dan Manfaat

Tujuan Tugas Akhir ini adalah untuk mengimplementasikan metode MVMCNN pada sistem identifikasi pengguna menggunakan biometrik *keystroke* untuk mengatasi kelemahan keamanan konvensional yang rentan terhadap peretasan. Tujuan spesifiknya adalah untuk melihat performansi identifikasi melalui penggunaan algoritma MVMCNN yang diterapkan pada Biomey *Keystroke Dataset* dari Universitas Telkom. Evaluasi performansi dilakukan dengan mengukur tingkat akurasi sistem menggunakan metrik *F1-score* sebagai parameter utama untuk menilai keberhasilan model.

1.4. Batasan Masalah

Dalam penelitian ini, terdapat beberapa batasan yang diterapkan untuk menyederhanakan ruang lingkup permasalahan. Pertama, jenis biometrik yang digunakan difokuskan pada implementasi sistem *keystroke* untuk identifikasi pengguna, bukan untuk verifikasi atau autentikasi. Pendekatan ini dipilih karena tujuan utama penelitian adalah melakukan klasifikasi pengguna secara unik berdasarkan pola pengetikan mereka. Kedua, *dataset* yang digunakan adalah Biomey *Keystroke Dataset*, yang terdiri dari data partisipan sebanyak 40 orang dari Universitas Telkom. Ketiga, evaluasi performa model dilakukan menggunakan metrik *F1-score* sebagai parameter utama. Terakhir, penelitian ini tidak membahas setiap fitur di dalam *dataset* secara mendalam, tetapi menggunakan keseluruhan fitur dasar *press* dan *release time* yang digunakan untuk menghasilkan lima fitur baru, yaitu UD (*Up-Down*), DD (*Down-Down*), DU (*Down-Up*), UU (*Up-Up*), dan *Duration*, yang merepresentasikan pola *keystroke* pengguna. Namun, penelitian ini tidak

melakukan observasi dan analisis untuk setiap fitur pada lima fitur *keystroke* tersebut.

1.5. Metode Penelitian

Penelitian ini dimulai dengan kajian literatur untuk memahami dasar-dasar *keystroke*, biometrik, dan algoritma klasifikasi MVMCNN. Tahap ini mencakup pengumpulan referensi ilmiah untuk membangun dasar teoretis yang kuat serta mengevaluasi pendekatan-pendekatan yang telah digunakan pada penelitian sebelumnya. Setelah itu, algoritma MVMCNN diterapkan pada Biomey *Keystroke Dataset*. *Dataset* ini diproses terlebih dahulu untuk memastikan kualitas dan konsistensinya, sebelum digunakan dalam model klasifikasi. Proses implementasi dilakukan dengan menggunakan bahasa pemrograman Python untuk membangun sistem identifikasi pengguna. Hasil dari sistem ini kemudian dievaluasi menggunakan *F1-score* untuk mengukur sejauh mana sistem dapat mengklasifikasikan data dengan akurat.

1.6. Jadwal Pelaksanaan

Jadwal pelaksanaan tugas akhir agar setiap tahapan dapat berjalan dengan baik dan terorganisir. Studi literatur dilakukan sepanjang enam bulan sebagai fondasi utama untuk memahami teori dan metode yang mendukung penelitian ini. Pengumpulan data selesai pada bulan pertama, sehingga fokus di bulan berikutnya dapat beralih ke tahap *pre-processing* model, yang berlangsung dari bulan kedua hingga bulan kelima, guna memastikan data siap digunakan dalam proses pembuatan model.

Tahap pembuatan dan implementasi model MVMCNN dilakukan paralel dengan *pre-processing* pada bulan ketiga hingga bulan kelima, diikuti oleh pengujian dan evaluasi mulai bulan ketiga hingga bulan keenam untuk

memastikan model yang dikembangkan dapat memberikan hasil yang sesuai. Sementara itu, penyusunan laporan tugas akhir mulai dilakukan sejak bulan kedua hingga bulan keenam, mendokumentasikan setiap proses dan hasil yang dicapai. Dengan jadwal ini, penelitian dapat berjalan lancar dan menghasilkan keluaran yang sesuai harapan.

Tabel 1.1 Jadwal Pelaksanaan Tugas Akhir.

No.	Deskripsi Tahapan	Bulan 1	Bulan 2	Bulan 3	Bulan 4	Bulan 5	Bulan 6
1	Studi Literatur						
2	Pengumpulan Data						
3	Pre-processing Model						
4	Pembuatan dan Implementasi Model MVMCNN						
5	Pengujian dan Evaluasi						
6	Penyusunan Laporan/Buku TA						