

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Autentikasi merupakan bagian dari sistem keamanan yang didasari oleh sesuatu yang diketahui atau dimiliki oleh pengguna (Bailey, Okolica, & Peterson, 2014). Contohnya adalah penggunaan *Common Access Card* (CAC) dan nomor pin atau username dan password (Matyáš & Říha, 2003). Autentikasi yang dilakukan hanya pada saat sesi masuk pengguna, masih bisa diambil alih dan informasi penting dicuri oleh penyerang. Namun ada pula autentikasi kontinu, yang dilakukan sepanjang sesi, dapat mencegah serangan tersebut (Balagani, 2015; Shen, Cai, & Guan, 2012). Bukan hanya barang atau *knowledge* yang bisa digunakan untuk autentikasi, fisiologi dan kebiasaan manusia dapat digunakan untuk autentikasi. Kedua hal tersebut termasuk ke dalam teknik biometrik.

Salah satu teknik autentikasi adalah teknik biometrik yang merupakan solusi menjanjikan dan studi yang penting dalam dunia sistem keamanan (Kodituwakku, 2015). Biometrik memiliki beberapa keunggulan dibandingkan cara tradisional, karena tidak dapat dicuri seperti barang (Prabhakar, Pankanti, & Jain, 2003). Sistem biometrik pada dasarnya adalah sistem pengenalan pola yang mengenali seseorang berdasarkan fitur vektor yang didapat dari karakteristik fisik atau kebiasaan dari orang yang memilikinya (Prabhakar et al., 2003). Salah satu teknologi biometrik semisal pengenalan sidik jari (Hrechak & McHugh, 1990) merupakan pengaplikasian biometrik dari fitur fisik manusia, selain itu teknik biometrik lainnya yang populer seperti gerakan tubuh dan kebiasaan penggunaan tetikus dan keyboard (Arshad, Wang, & Chen, 2015; Bailey et al., 2014; Chuda, Kratky, & Tvarozek, 2015; Karim & Heickal, 2017) merupakan biometrik kebiasaan (*behavioral biometric*).

Sistem pengenalan *behavioral biometric* memiliki kelebihan dibandingkan fisiologis karena tidak memerlukan sensor khusus untuk

menangkap masukan. Perangkat yang biasa atau wajib dalam sebuah sistem komputer sudah cukup untuk mengaplikasikan biometrik *behavioral*. Sistem ini mengenali kebiasaan seseorang yang merupakan hasil dari faktor psikologi dan fisiologi masing-masing (Bailey et al., 2014). Psikologi dan fisiologi yang dimaksud adalah kebiasaan pengguna dalam mengoperasikan komputer dan bentuk tubuh yang mengoperasikannya contohnya dengan kebiasaan menggunakan papan ketik dan tetikus. Gerakan Tetikus menjadi pendekatan yang cukup menjanjikan. Pendekatan ini memanfaatkan fitur-fitur kebiasaan yang intrinsik dalam penggunaan tetikus untuk mendeteksi penyusup (Shen et al., 2012).

Metode behavioral yang telah diteliti meliputi *keystroke dynamics* (Brown & Rogers, 1993; Gunetti & Picardi, 2005; Joyce & Gupta, 1990; Marsters, 2009; Monroe & Rubin, 1997), *mouse dynamics* (Ahmed & Traore, 2007; Shen, Cai, Guan, Du, & Masion, 2013; Zheng, Paloski, & Wang, 2011), pengenalan suara (Bhattacharyya, Ranjan, a, & Choi, 2009), *signature verification* (Bhattacharyya et al., 2009) dan analisis penggunaan antarmuka pengguna grafis (Gamboa & Ana, 2004; Imsand, 2008). Dikarenakan keragaman dari tubuh dan pikiran manusia, pengadaptasian biometrik tipe ini tertinggal dari biometrik fisik manusia. Walaupun begitu, biometrik yang menggunakan metode *keystroke dynamic*, *mouse dynamic*, dan interaksi antarmuka tidak memerlukan tambahan perangkat keras (sebagai sensor) (Bailey et al., 2014)

Pada perkembangannya, sistem autentikasi biometrik tetikus dihadapkan pada permasalahan pemilihan fitur dan pengolahan fitur pergerakan. Banyak penelitian biometrik tetikus berfokus pada ekstraksi fitur dan skenario perekaman tetapi sedikit yang meneliti tentang pengaruh metode reduksi dimensi dan fungsi kernel pada saat klasifikasi (Bailey et al., 2014; Villani et al., 2006). Shen et al. melakukan eksperimen dengan definisi fitur sendiri dan menggunakan klasifikasi *oneclass support vector machine* dengan alasan bahwa autentikasi biometrik tetikus masuk ke dalam *novelty detection* yang tidak memiliki kelas negatif di dunia nyata.

Autentikasi biometrik tetikus masuk ke dalam *novelty detection*, dimana sistem bertujuan untuk menemukan anomali dari data masukan yang baru. Pada situasi tersebut dataset yang digunakan untuk *training* lebih banyak kelas positif dibanding negatif atau bahkan tidak ada dataset kelas negatif karena ketersediaan data hanya kondisi positif (Azkiya, Indriani, & Chandra, 2017). Maka dari itu, banyak klasifikasi biometrik menggunakan *oneclass support vector machine*. Metode klasifikasi ini menerapkan klasifikasi 2 kelas dengan hanya 1 kelas dataset. *Oneclass support vector machine* akan mencari *hyperplane* sebagai pembatas kelompok data normal dan akan menganggap data di luar *hyperplane* sebagai anomali (Schölkopf, Williamson, Smola, Shawe-Taylor, & Platt, 2000).

Pada prinsipnya *hyperplane* pada *support vector machine* adalah garis lurus berupa fungsi linear. Namun, *hyperplane* linear tidak bisa digunakan dalam kasus *novelty detection*. Hal ini dikarenakan garis pembatas harus berupa kurva tertutup dimana ia membungkus kelompok data normal dalam area tertentu dengan asumsi data anomali berada diluar area tersebut. Selain itu, banyaknya fitur dalam biometrik tetikus yang sulit untuk dipisahkan secara linear. Maka dari itu diperlukan sebuah fungsi untuk mengubah *hyperplane* linear menjadi nonlinear yaitu dengan fungsi kernel. Fungsi kernel atau *kernel trick* mengubah garis linear menjadi nonlinear dengan memasukkan titik hasil dari *hyperplane* ke fungsi kernel maka *hyperplane* akan berubah menjadi nonlinear. Kernel semisal *radial base function* akan mengubah *hyperplane* mendekati bentuk lingkaran, sementara kernel polinomial akan mengubah *hyperplane* ke bentuk polinomial.

Secara umum, dalam proses klasifikasi tidak semua fitur dalam dataset memiliki proporsi yang sama. Fitur-fitur dengan nilai variansi yang kecil memiliki proporsi yang kecil sehingga tidak terlalu berpengaruh dalam klasifikasi. Sementara itu bila seluruh fitur dimasukkan ke dalam klasifikasi maka akan memberati komputasi. Digunakanlah reduksi dimensi dengan *principal component analysis*. Metode akan mereduksi data dengan

mentransformasi data ke dalam ruang dimensi baru (Chandra, Novianty, Jati, Elektro, & Telkom, 2015), dengan begitu akan didapatkan fitur baru yang terurut dari proporsi yang paling besar ke yang paling kecil (Wardoyo, Wiryadinata, Sagita, & Presensi, 2014). Fitur-fitur dengan proporsi besar inilah yang akan diambil untuk klasifikasi karena memiliki korelasi kuat, secara bersamaan mereduksi dimensi data dengan membuang fitur dengan proporsi kecil karena komponen tersebut tidak terlalu berkorelasi dengan data lainnya (Pramudhito, Purwanto, & Novianty, 2017).

Shen pada 2013 melakukan reduksi dimensi menggunakan *kernel principal component analysis*. Penggunaan fungsi kernel dalam *principal component analysis* memberikan beban komputasi yang berat (Hoffmann, 2007). Hal ini sangat tidak diinginkan dalam sistem autentikasi yang mana kecepatan merupakan hal paling dirasakan oleh pengguna. Selain itu pemilihan fungsi kernel yang tidak tepat dapat memperlambat proses autentikasi dimana kecepatan merupakan faktor penting selain akurasi dalam sistem autentikasi.

Penelitian sebelumnya oleh Shen tidak meneliti pengaruh jumlah *principal component* reduksi dimensi terhadap akurasi sistem. Sistem diberi kebebasan untuk menentukan jumlah *principal component* yang digunakan. Hal ini memunculkan pertanyaan bagaimana pengaruh jumlah *principal component* terhadap akurasi sistem autentikasi biometrik tetikus dan konfigurasinya. Maka dari itu diperlukan analisis jumlah *principal component* pada sistem ini dan berapa batas kelayakan jumlah *principal component* yang dipakai.

Berdasarkan beberapa penjelasan yang telah disampaikan, penulis ingin membangun dan menganalisa prototipe sistem autentikasi menggunakan *behavioral* biometrik tetikus dengan pendekatan ekstraksi fitur holistic dan prosedural (Shen et al., 2013) dengan menerapkan *principal component analysis* (PCA) sebagai metode reduksi dimensi dan *kernelized oneclass support vector machine* (SVM) sebagai *classifier*.

## 1.2. Rumusan Masalah

Berdasarkan permasalahan yang muncul di latar belakang, maka beberapa masalah yang ingin diselesaikan yaitu:

1. Bagaimana konfigurasi reduksi dimensi dengan PCA pada autentikasi biometrik tetikus?
2. Bagaimana tingkat akurasi sistem autentikasi biometrik tetikus menggunakan PCA?
3. Bagaimana kecepatan pengenalan behavioral biometric dengan metode PCA dan kernelized oneclass SVM dibandingkan dengan kernel PCA?

## 1.3. Tujuan

Penelitian ini bertujuan untuk membangun prototipe sistem autentikasi pergerakan tetikus menggunakan *principal component analysis* dan *oneclass support vector machine*. Berikut adalah beberapa sub-tujuan yang ingin dicapai dalam penelitian ini :

1. Menganalisis pengaruh reduksi dimensi dari metode *principal component analysis* dengan berbagai jumlah komponen hasil reduksi.
2. Menganalisis tingkat akurasi sistem autentikasi biometrik tetikus dengan metode PCA sebagai pereduksi dimensi.
3. Melakukan eksperimen kecepatan autentikasi biometrik tetikus dengan metode eksperimen pada lingkungan yang sama.

## 1.4. Batasan Masalah

Terdapat beberapa batasan masalah dalam penelitian ini sebagai berikut:

1. Data hasil rekaman mahasiswa departemen ilmu komputer yang dilakukan oleh penulis.
2. Menggunakan konsep yang dilakukan oleh Shen et al. Fitur pergerakan tetikus terbagi menjadi 6 jenis yang dikategorikan dalam fitur holistik dan prosedural.

3. Model dibangun dalam lingkungan terkontrol. Yaitu setiap *task* perekaman gerakan tetikus dibuat sama dan statis untuk semua subjek dan resolusi layar yang digunakan sama.
4. Aplikasi perangkat lunak dibangun dengan bahasa pemrograman R

### **1.5. Sistematika Penulisan**

Berikut ini adalah sistematika penulisan yang dilakukan dalam menyusun skripsi :

#### **BAB I PENDAHULUAN**

Bab ini berisikan masalah yang diangkat dalam penelitian meliputi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, dan sistematika penulisan.

#### **BAB II KAJIAN PUSTAKA**

Berisi tentang kajian teori beserta *literature map* yang digunakan dalam penelitian meliputi teori tentang autentikasi, biometrik, *support vector machine*, Neural network, tetikus, *principal component analysis*, fungsi kernel, dan pengenalan gerakan tetikus.

#### **BAB III METODOLOGI PENELITIAN**

Berisikan deskripsi langkah-langkah dan metode yang digunakan dalam melakukan penelitian meliputi data, pra-proses, *dimensionality reduction*, *classifier*, lingkungan sistem, rancangan eksperimen, dan tahapan evaluasi.

#### **BAB IV METODOLOGI PENELITIAN**

BAB IV menjabarkan hasil penelitian yang telah dilakukan dan analisis kualitas sistem yang telah dibuat, dengan mengukur tingkat akurasi, tingkat kecepatan autentikasi, dan perbandingan dengan penelitian terkait. Semua pertanyaan mengenai masalah yang diangkat dalam tema skripsi dibahas disini. Bab ini terdiri dari Analisis dan hasil eksperimen.

## **BAB V KESIMPULAN DAN SARAN**

BAB V menjelaskan mengenai penarikan kesimpulan dari hasil penelitian, serta memaparkan beberapa saran untuk penelitian selanjutnya. Bab ini berisi kesimpulan dari hasil eksperimen dan saran untuk pengembangan penelitian selanjutnya.