

BAB I

PENDAHULUAN

A. Latar Belakang

Pada era modern saat ini penggunaan teknologi *biometrik* seperti *fingerprint* dalam pengenalan individu semakin penting dalam keamanan dan pengendalian akses. Sidik jari merupakan elemen krusial dalam pengenalan individu, karena kemampuannya untuk memastikan identitas dengan akurasi tinggi. Oleh karena itu, akurasi dalam pencocokan sidik jari sangat diperlukan untuk mengurangi kesalahan yang dapat berdampak signifikan (Growth et al., 2022). Dalam penelitian ini klasifikasi yang akurat dalam pengenalan *fingerprint* dapat dicapai dengan menggunakan algoritma *K-Nearest Neighbor* (KNN), yang merupakan salah satu algoritma pembelajaran mesin (*machine learning*) yang paling sederhana namun efektif. Namun, masalah dalam penerapan KNN adalah menentukan nilai K yang optimal. Nilai K yang kurang optimal dapat menghasilkan hasil klasifikasi yang tidak akurat.

K-Nearest Neighbor (KNN) merupakan algoritma yang digunakan untuk mengklasifikasikan setiap data yang tidak memiliki label dengan menggunakan mayoritas label dari k -tetangga terdekat pada sekumpulan data latih (Dubey & Pudi, 2013). Oleh karena itu, kinerja KNN sangat bergantung pada metrik jarak yang digunakan untuk mencari tetangga terdekatnya. Dalam algoritma KNN tetangga terdekat dihitung sesuai dengan nilai K yang menentukan jumlah tetangga terdekat untuk dipertimbangkan termasuk ke dalam kelas apa pada data latih (Amra, 2017). Sebagian besar pengklasifikasi KNN menggunakan metrik jarak *Manhattan* untuk mengukur perbedaan antara target yang direpresentasikan sebagai *input* dengan data latih.

Algoritma KNN dipilih pada kasus ini karena memiliki beberapa kelebihan yang pertama yaitu Probabilitas Kesalahan Rendah, algoritma *KNN* dikenal dengan probabilitas kesalahan yang rendah, menjadikannya pilihan yang andal untuk tugas klasifikasi, terutama saat menangani kumpulan data yang kompleks di mana akurasi

sangat penting (Yahia & Ibrahim, 2004). Kemudian yang kedua adalah Robustness, algoritma ini dapat diperkuat dengan teknik statistik dan perbaiki dengan algoritma lain seperti algoritma genetika untuk menghilangkan bias pada nilai K . Robustness ini penting untuk memastikan kinerja algoritma tetap optimal dalam berbagai skenario dan kumpulan data (Yahia & Ibrahim, 2004). Meskipun memiliki banyak kelebihan, algoritma KNN tetap memiliki beberapa kekurangan salah satu kekurangan atau kelemahan yang cukup jelas dari algoritma KNN adalah kesulitan dalam menentukan nilai K awal yang optimal. Maka dari itu pemilihan parameter K dari KNN memiliki pengaruh tertentu terhadap kinerja klasifikasi KNN (Fei, 2019). Jika nilai K terlalu kecil maka dapat memberikan hasil klasifikasi yang mudah terpengaruh oleh noise. Di sisi lain, jika nilai K terlalu tinggi akan mengurangi efek noise pada klasifikasi, namun membuat batasan antara setiap klasifikasi menjadi lebih kabur. Nilai K yang optimal dapat dipilih dengan melakukan optimasi pada parameter K tersebut. Tingkat keakurasian pada klasifikasi ini tergantung pada pemilihan nilai K terbaik, untuk itulah nilai K pada algoritma KNN perlu dilakukan optimasi. Pengoptimasian ini dilakukan dengan menggunakan *Particle Swarm Optimization* (PSO), algoritma ini memiliki manfaat dalam menghasilkan nilai K yang optimal yang selanjutnya dapat diaplikasikan pada algoritma KNN.

Particle Swarm Optimization merupakan algoritma pencarian berbasis populasi dan inialisasi populasi secara acak yang disebut dengan partikel atau individu (Rakhman, 2017). Penggunaan PSO untuk optimasi pada penelitian ini dikarenakan proses pengoptimalan yang lebih cepat, lebih sedikit parameter yang digunakan, dan keseimbangan yang terjaga (Ayun *et al.*, 2022). PSO dipilih dengan beberapa alasan antara lain memiliki konsep yang sederhana, dengan konvergensi yang relatif cepat, mudah diimplementasi dalam berbagai bidang guna menemukan pemecahan masalah terkait dengan optimasi (Arsi *et al.*, 2021).

Dengan demikian, penelitian ini bertujuan untuk meningkatkan akurasi klasifikasi *fingerprint* dengan mengoptimalkan nilai K pada algoritma KNN menggunakan PSO. Penerapan PSO untuk optimasi pada KNN diharapkan

mendapatkan hasil akurasi yang baik serta dapat dicapainya penyelesaian klasifikasi pada *fingerprint* di penelitian ini.

B. Perumusan Masalah

Algoritma KNN memiliki kelemahan pada penentuan nilai K awal yang menjadikan diperlukan adanya optimasi untuk menentukan nilai K awal terbaik. Oleh karena itu penelitian ini dilakukan untuk mencari nilai K awal yang optimal dengan menggunakan PSO pada data uji *fingerprint*.

C. Batasan Masalah

Batasan masalah pada penelitian ini dipaparkan sebagai berikut.

1. *Dataset* yang digunakan merupakan data yang didapat dari *repository Kaggle*.
2. Dalam penelitian ini, hasil akhir kinerja algoritma akan dievaluasi menggunakan *Confusion Matrix* yang akan meliputi nilai akurasi (*accuracy*), presisi (*precision*), dan *recall*.

D. Tujuan Penelitian

Penelitian ini dilakukan bertujuan untuk mengoptimalkan algoritma KNN dalam klasifikasi data *fingerprint* dengan menggunakan *Particle Swarm Optimization* (PSO) untuk menentukan nilai K awal yang optimal. Dalam hal ini, penelitian akan memfokuskan pada pengembangan teknik optimasi menggunakan PSO untuk menyelesaikan permasalahan nilai K awal pada algoritma KNN.

E. Manfaat Penelitian

Manfaat yang diharapkan dalam penelitian ini adalah sebagai berikut.

1. Dapat mengetahui cara kerja *Particle Swarm Optimization* dalam melakukan optimasi.
2. Dapat mengetahui efektivitas klasifikasi terhadap data uji *fingerprint* dengan *K-Nearest Neighbor* yang di optimasi dengan *Particle Swarm Optimization*.
3. Dapat menjadi acuan bagi penelitian selanjutnya yang menyangkut tentang optimasi dengan menggunakan *Particle Swarm Optimization* ataupun klasifikasi data dengan menggunakan *K-Nearest Neighbor*.

4. Dapat memberikan kontribusi dalam meningkatkan akurasi dan efisiensi klasifikasi data *fingerprint* yang digunakan pada sistem pengamanan atau identifikasi personal.

