

---

## 1. Pendahuluan

### Latar Belakang

Latar belakang penelitian ini mengacu pada pemilihan Twitter sebagai subjek utama analisis. Twitter merupakan salah satu platform media sosial yang telah mencapai popularitas yang besar dalam era digital saat ini. Twitter memungkinkan penggunanya untuk menulis dan membagikan pendapat mereka melalui tulisan pribadi atau konten yang dibagikan oleh orang lain. Menurut data dari Statistika, pada kuartal II-2021, Twitter memiliki 206 juta pengguna aktif harian di seluruh dunia.[1]. Selain itu pada media sosial Twitter juga terdapat fitur *Retweet*. *Retweet* merupakan fitur yang memposting ulang sebuah *tweet* baik *tweet* diri sendiri maupun *tweet* orang lain yang mengartikan kita juga menyukai atau setuju dengan *tweet* tersebut. Postingan/*tweet* yang diretweet biasanya bertujuan untuk mempublikasikan lebih luas lagi agar *tweet* tersebut dapat dijangkau oleh lebih banyak orang.

Twitter memberikan keterbatasan dalam penulisan, dengan batasan maksimal sebanyak 140 karakter dalam setiap posting (*tweet*). Pengguna Twitter, baik yang terdaftar maupun yang bukan pengguna, dapat memantau perkembangan topik tertentu dengan mengikuti tagar (*hashtag*) yang ditandai dengan simbol tanda pagar (#). Tanda pagar ini memungkinkan individu untuk mengikuti serta berpartisipasi dalam diskusi seputar topik tertentu yang menarik perhatian.[2]. Maka dari itu, platform twitter memiliki fitur di mana seseorang dapat menemukan informasi dari suatu kata kunci yang sedang banyak dibicarakan oleh banyak pengguna twitter.

Dalam penelitian ini, peneliti mengembangkan model dengan menggunakan 16.000 data *tweet* yang diperoleh melalui perpustakaan Tweepy. Kemudian, model ini menggunakan dua jenis fitur, yaitu fitur yang berkaitan dengan pengguna dan fitur yang berkaitan dengan konten. Data *tweet* yang digunakan dalam penelitian ini ditulis dalam bahasa Indonesia, dan tujuan penelitian ini adalah membangun model yang dapat memprediksi dan mengukur keakuratan apakah suatu konten akan di-*retweet* atau tidak.

Model prediksi *retweet* ini memanfaatkan Jaringan Saraf Tiruan yang telah dioptimalkan menggunakan Particle Swarm Optimization (PSO). PSO sendiri adalah sebuah algoritma berbasis populasi yang mengoptimalkan individu-individu dalam proses pencarian. Dalam PSO, populasi disebut sebagai "swarm" dan individu-individu disebut sebagai "particle." Setiap partikel bergerak dengan kecepatan yang disesuaikan berdasarkan informasi dari lingkungan pencarian, dan mereka menyimpan catatan dari posisi terbaik yang pernah mereka capai. [3].

### Topik dan Batasannya

Topik pada penelitian ini adalah melakukan pengujian model jaringan syaraf tiruan yang dioptimasi *Particle Swarm Optimization*. Penelitian ini memiliki batasan masalah yaitu dataset yang digunakan merupakan *dataset* Bahasa Indonesia dengan menggunakan kata kunci *blackpink*. Dataset tersebut terdiri dari 16000 data *tweet* terhitung dari bulan Maret-Juni 2023.

### Tujuan

Tujuan penelitian ini adalah untuk melihat performansi dari model jaringan syaraf tiruan yang dioptimasi *Particle Swarm Optimization* untuk melakukan prediksi *retweet*. Skenario pengujian yang dilakukan pada penelitian yaitu ada dua skenario, skenario pertama menggunakan *Random Under Sampling* dan skenario kedua menggunakan *Random Over Sampling*. Langkah selanjutnya adalah menghitung akurasi dan F1-score dari model JST-PSO dan melakukan perbandingan model JST saja dengan JST-PSO.

### Organisasi Tulisan

Setelah bagian pendahuluan dalam Bab 1, tahapan berikutnya pada Bab 2 akan mengulas studi literatur yang relevan dengan kegiatan penelitian yang direncanakan. Pada Bab 3, akan dibahas teori dan sistem yang diimplementasikan dalam penelitian ini. Bab 4 akan membahas hasil analisis dari penelitian yang telah dilaksanakan, diikuti oleh penjelasan kesimpulan yang diperoleh berdasarkan temuan penelitian pada Bab 4. Kesimpulan tersebut akan dijelaskan lebih lanjut dalam Bab 5.