

1. Pendahuluan

1.1 Latar Belakang

Di era digital saat ini, sosial media telah menjadi alat komunikasi utama dan sarana pertukaran informasi bagi miliaran orang di seluruh dunia [1]. Platform seperti Instagram dengan 2 miliar pengguna, Twitter dengan 666 juta pengguna, dan YouTube dengan hampir 2,5 miliar pengguna, tidak hanya berfungsi sebagai saluran interaksi sosial tetapi juga memainkan peran penting dalam bisnis, politik, dan pendidikan [2]. Namun, popularitas ini juga membawa tantangan baru, terutama terkait keaslian akun. Akun ganda, yang sering digunakan untuk tujuan manipulatif atau berbahaya, telah menjadi masalah yang semakin menonjol [1]. Hal ini menimbulkan kekhawatiran terkait integritas dan keamanan online, membuat deteksi akun ganda menjadi kebutuhan yang mendesak [3].

Masalah akun ganda di sosial media bukan hanya tentang jumlah, tetapi juga dampaknya. Akun-akun tersebut dapat digunakan untuk memanipulasi diskusi publik, menyebarkan informasi palsu, atau bahkan melakukan penipuan. Di Instagram, misalnya, akun ganda dapat digunakan untuk meningkatkan popularitas seseorang atau produk secara artifisial [4]. Di Twitter, akun ganda dapat mempengaruhi narasi politik [5]. Dalam pemilihan calon presiden, akun ganda sering digunakan untuk memperkuat pesan kampanye, menyebarkan narasi yang menguntungkan salah satu pihak, atau bahkan menyerang lawan politik dengan informasi yang menyesatkan. Sementara itu, di YouTube, akun ganda dapat memanipulasi statistik penonton dan interaksi [6]. Hal ini dapat mempengaruhi persepsi publik terhadap popularitas dan dukungan terhadap kandidat tertentu.

Namun, mendeteksi akun ganda merupakan tantangan yang kompleks. Dari penelitian sebelumnya, Kumar et al. [3] telah meneliti kloning profil di Instagram dengan menggunakan *similarity index*. F. S. Alsubaei [7] telah menerapkan metode *random forest* untuk mendeteksi akun palsu di Twitter. Sementara itu, H. Oh [8] menggunakan *cascaded ensemble machine learning model* untuk mendeteksi spam di komentar YouTube. Akun-akun ini sering kali sulit dibedakan dengan pengguna asli karena meniru pola perilaku manusia [9].

Penggunaan Apify untuk *web scraping* adalah kunci dalam mengumpulkan data yang diperlukan dari Instagram dan Twitter untuk analisis [10]. Selain itu, penggunaan YouTube API v3 menyediakan data yang diperlukan dari YouTube [11]. *Deep Learning* dapat memproses dan menganalisis dataset yang besar dan kompleks, memberikan cara yang inovatif dan efisien untuk mengatasi masalah ini [12]. Dengan menggunakan model *Recurrent Neural Networks* (RNN) untuk data tekstual [13][14][15], *Deep Learning* dapat mengekstrak fitur-fitur penting dan mengidentifikasi pola-pola yang tidak terlihat dengan metode konvensional. Keunggulan dari *Deep Learning* terletak pada kapasitasnya untuk belajar dari data yang sangat besar, memungkinkan identifikasi yang lebih akurat dari perilaku akun ganda dan otomatisasi proses deteksi [13].

Kebaruan dari penelitian ini terletak pada jenis dataset yang digunakan, yang bersumber dari Apify dan YouTube API v3. Berdasarkan penelitian-penelitian sebelumnya, penelitian ini mengidentifikasi kriteria-kriteria yang dapat menjadi acuan untuk menentukan apakah sebuah akun terduplikasi. Selain itu, penerapan *deep learning*, khususnya *Recurrent Neural Networks* (RNN), memperkenalkan dimensi baru dalam pengembangan metode pendeteksian akun ganda, yang menjanjikan kemajuan yang signifikan dalam bidang ini. Pendekatan ini bertujuan untuk meningkatkan akurasi dan efisiensi pendeteksian akun ganda dengan memproses dan menganalisis dataset yang besar dan kompleks, memberikan solusi inovatif yang tidak tercakup dalam referensi utama.

1.2 Topik dan Batasannya

Penelitian ini berfokus pada deteksi dan klasifikasi akun ganda di media sosial, khususnya di platform seperti Instagram, Twitter, dan YouTube. Ketiga platform ini dipilih karena sering digunakan untuk memanipulasi diskusi publik, menyebarkan informasi palsu, atau bahkan melakukan penipuan. Data dikumpulkan melalui Apify untuk *web scraping* dan YouTube API v3, yang menghasilkan dataset yang besar dan kompleks. Dengan memproses dan menganalisis data tersebut, *Recurrent Neural Networks* (RNN) dapat mengekstraksi fitur-fitur penting dan mengidentifikasi pola-pola yang tidak terdeteksi oleh metode konvensional. Batasan penelitian ini terletak pada penggunaan dataset dari Apify dan YouTube API v3, yang mungkin memiliki keterbatasan dalam representasi data. Meskipun demikian, penelitian ini menawarkan pendekatan baru dalam deteksi akun ganda menggunakan *Deep Learning*, yang lebih akurat dan efisien dibandingkan dengan pendekatan sebelumnya.

1.3 Tujuan

Tujuan penelitian ini adalah untuk mencari karakteristik khusus yang membedakan akun ganda dari akun asli pada platform sosial media seperti Instagram, Twitter, dan YouTube, serta menguji performansi algoritma *Deep Learning*, khususnya *Recurrent Neural Networks* (RNN), dalam mendeteksi akun ganda pada platform sosial media Instagram, Twitter, dan Youtube.

1.4 Organisasi Tulisan

Struktur penelitian ini terdiri dari lima bagian yang memberikan gambaran yang komprehensif tentang keseluruhan studi. Bagian I memperkenalkan latar belakang, topik dan batasannya, tujuan dari penelitian ini, serta organisasi tulisan. Bagian II menyajikan tinjauan literatur yang mendalam mengenai karya-karya terkait, memberikan konteks yang lebih luas serta landasan teoretis yang mendukung penelitian ini. Selanjutnya, Bagian III membahas desain sistem yang digunakan dalam penelitian, termasuk metode, alat, dan teknik yang diterapkan untuk mencapai tujuan penelitian. Bagian IV menyajikan hasil dan rincian eksperimen yang dilakukan, mencakup analisis data, evaluasi performa algoritma, dan interpretasi dari temuan yang diperoleh. Terakhir, Bagian V menyimpulkan penelitian ini dengan rangkuman temuan utama, diskusi mengenai implikasi dari hasil penelitian, serta saran untuk penelitian selanjutnya yang dapat memperluas atau memperdalam studi di bidang ini.