

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi yang berkembang dengan cepat telah memberikan manfaat yang dirasakan secara luas oleh masyarakat di seluruh dunia. Kebutuhan masyarakat dapat terpenuhi dengan lebih mudah berkat adanya teknologi informasi. Salah satu bentuk perkembangan teknologi informasi yang signifikan adalah munculnya media online. Media online dianggap sebagai sumber utama informasi yang berdampak besar pada stabilitas kehidupan sosial masyarakat [1]. Salah satu bentuknya adalah media sosial yang saat ini menjadi sebuah alat yang sangat efektif dalam menyebarkan informasi kepada masyarakat umum. Media sosial adalah bagian integral dari kemajuan teknologi informasi yang memungkinkan individu untuk mengekspresikan ide dan pemikiran mereka dengan luas. Selain itu, media sosial juga sering dijadikan wadah untuk mengekspresikan ide-ide kreatif baru kepada masyarakat, menghasilkan inovasi dan gagasan baru setiap harinya [2].

Kemampuan media online untuk mencapai audiens yang luas melalui internet memungkinkan kinerjanya menjadi lebih efisien. Karena kurangnya aturan yang ketat dalam pengaturan penulisan informasi, menyebarkan informasi yang sulit diverifikasi menjadi sangat mudah di *platform online*. Hal ini berarti siapa pun dapat mengakses media *online* dan dengan leluasa menyebarkan informasi tanpa melewati proses seleksi. Fenomena ini menunjukkan bahwa informasi dapat disampaikan secara anonim atau tanpa sumber yang jelas. Dampaknya, informasi yang disebar seringkali tidak sesuai dengan kenyataan dan fakta yang terjadi, sehingga sering dianggap sebagai hoaks dan berpotensi memunculkan ujaran kebencian [3].

Setiap harinya, portal berita dapat memuat banyak berita hoaks yang memicu kontroversi, memunculkan pendapat berbeda, dan menghasilkan konflik serta debat yang sering terjadi di kolom komentar unggahan tersebut. Penyebaran berita

hoaks tanpa bertanggung jawab terhadap kebenaran fakta sesuai dengan pendapat Husin, berita adalah laporan harian mengenai fakta atau ide terbaru yang dipilih oleh redaksi untuk dipublikasikan dan menarik perhatian pembaca karena relevansinya atau dampaknya. Berita sering kali mencakup aspek *human interest* seperti unsur humor, emosi, dan ketegangan, yang mempengaruhi cara informasi disampaikan dan diterima [4]. Berita hoaks adalah bentuk penipuan yang disebarkan melalui pemberitaan atau informasi palsu yang tersebar di masyarakat. Berita hoaks memengaruhi pembaca dengan menyajikan informasi yang tidak akurat sehingga mengarahkan mereka untuk bertindak tidak sesuai dengan kenyataan. Sehingga pembaca mudah sekali untuk digiring opininya dan mempercayai informasi yang tidak sesuai dengan faktanya, terutama dalam konteks isu politik. Pertentangan meningkat dalam politik tidak hanya menghalangi keputusan yang efektif, tetapi juga merusak hubungan sosial dan mempersulit pencapaian kesepakatan dalam masyarakat. Situasi ini dapat memperbesar perpecahan dan menghambat kemajuan dalam proses pengambilan keputusan yang mengharuskan kerjasama antarberbagai pandangan [5].

Menurut Kementerian Komunikasi dan Informatika, teridentifikasi sebanyak 425 isu hoaks selama triwulan pertama tahun 2023. Angka ini mengalami peningkatan dibandingkan dengan triwulan pertama tahun 2022 yang mencatat 399 isu hoaks. Saat ini, berbagai *platform* seperti aplikasi chat seperti WhatsApp, LINE, Instagram, X, dan lainnya menjadi sumber penyebaran hoaks. Fenomena ini menjadi topik menarik untuk diteliti mengingat berbagai peristiwa baru-baru ini yang telah menimbulkan kekhawatiran di tengah masyarakat. Menurut Budi Arie Setiadi selaku Menteri Komunikasi dan Informatika menyatakan bahwa pada tahun 2022 tercatat hanya 10 hoaks terkait Pemilu. Namun, dari Januari 2023 hingga 26 Oktober 2023, jumlah isu hoaks Pemilu telah mencapai 91. Ini menandakan peningkatan hampir sepuluh kali lipat jumlah isu hoaks dibandingkan dengan tahun sebelumnya [6].

Saat ini masih belum ada penanganan yang cukup baik dalam mengatasi penyebaran hoaks. Karena komentar yang terus-menerus diunggah oleh

masyarakat, tidak dapat diatur oleh siapapun bahkan pemerintahan. Hal ini memungkinkan timbulnya keributan jika hanya dipantau secara *manual* melalui media sosial terkait penyebarannya berita hoaks. Sehingga diperlukan analisa untuk mengidentifikasi dan mengelompokkan informasi palsu yang menggunakan pendekatan canggih dalam analisis data. Algoritma-algoritma seperti K-Means dan DBSCAN memiliki relevansi yang signifikan dalam konteks ini karena kemampuan mereka dalam pengelompokan data.

K-Means adalah salah satu algoritma klustering yang sangat berguna dalam mengidentifikasi pola dalam data teks. Algoritma ini dikenal karena kemampuan komputasinya yang cepat dan kemudahan dalam interpretasi pola kluster. Dalam konteks mengatasi hoaks, K-Means dapat membantu dalam mengelompokkan konten teks yang serupa atau memiliki pola yang mirip. Hal ini memungkinkan untuk mengidentifikasi kelompok-kelompok tertentu dari konten yang memiliki karakteristik yang serupa, memudahkan dalam pengelompokan informasi palsu yang mungkin menyebar.

Di sisi lain, DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) menggunakan pendekatan berbasis kepadatan data. Algoritma ini mampu memisahkan data yang tidak beraturan atau *noise* dari kelompok data yang valid berdasarkan kepadatan datanya. DBSCAN menghasilkan tiga jenis status untuk setiap titik data, seperti *core*, *border*, dan *noise* [7]. Titik-titik inti (*core points*) memiliki kepadatan yang cukup untuk membentuk kluster, sedangkan titik-titik batas (*border points*) berdekatan dengan kluster tapi tidak memiliki kepadatan yang cukup untuk dianggap sebagai inti, dan titik-titik *noise* tidak termasuk ke dalam kluster apa pun karena kepadatannya yang rendah.

Dalam konteks informasi palsu, DBSCAN dapat membantu dalam mengidentifikasi kelompok-kelompok atau pola yang tidak hanya berdasarkan kesamaan teks, tetapi juga kepadatan informasi yang tersebar. Ini memungkinkan untuk mengelompokkan data yang mungkin memiliki pola yang tidak teratur atau tidak seragam secara lebih efisien, membantu dalam mendeteksi atau memfilter informasi palsu yang tersebar di antara data yang valid. Dengan kombinasi

algoritma-algoritma ini, analisis data dapat memainkan peran krusial dalam upaya menghadapi dan menanggulangi penyebaran informasi palsu di era digital saat ini. Dengan mengintegrasikan kedua algoritma ini dalam analisis teks, dapat membuka jalan untuk memahami dan mengenali pola-pola yang menandai keberadaan berita palsu dengan lebih baik.

Berdasarkan penelitian sebelumnya yang telah dijalankan oleh Yashoda dan Jatidenkurmar di tahun 2022 dengan judul *Unsupervised Misinformation Detection Model using Incremental KMeans Algorithm*. Penelitian ini memperkenalkan model deteksi misinformasi tanpa pengawasan yang adaptif terhadap data baru *real-time* dengan algoritma K-Means inkremental. Model ini berhasil dengan skor kemurnian 0,92% dan skor siluet rata-rata 0,57%. Analisis menunjukkan kluster C2 lebih banyak berisi URL palsu, sementara C1 lebih banyak URL benar. C2 juga memiliki lebih banyak penerbit buruk, menandakan kecenderungan menyebarkan misinformasi, terutama terkait dengan Iran. Penelitian ini memberikan kontribusi pada pengembangan model deteksi adaptif dan wawasan tentang reputasi, penerbit, serta penyebaran misinformasi [8].

Selain itu, pada penelitian yang berjudul *Clustering Content Types and User Motivation using DBSCAN on Twitter*, penelitian ini menggunakan DBSCAN pada 49.519 *tweet* Bahasa Indonesia tentang Anies Baswedan sebagai Calon Presiden 2024. *Clustering* dilakukan dengan variasi parameter, menunjukkan variasi jumlah kluster tergantung pada parameter. Analisis kluster membantu mengidentifikasi jenis konten dan motivasi pengguna, memberikan wawasan tentang sentimen politik di Twitter. Penelitian ini berperan dalam pengembangan metode analisis teks dan kluster untuk data media sosial serta memperdalam pemahaman tentang pandangan pengguna Twitter terhadap Anies Baswedan sebagai Calon Presiden 2024 [9].

Setelah menguraikan fenomena serta dampak negatif yang muncul seiring penyebaran berita palsu atau hoaks menjelang Pemilihan Presiden di tahun 2024, penelitian ini dilakukan untuk mendeteksi konten hoaks tersebut. Penelitian ini menggunakan dataset hasil pengumpulan data melalui *platform X* yang diambil

dari 1 Januari hingga 21 Mei 2023, serta menggunakan algoritma *clustering* K-Means dan DBSCAN.

1.2 Rumusan Masalah

Dari konteks pembahasan yang telah diuraikan, berikut beberapa perumusan masalah yang mungkin terdapat dalam penelitian ini:, yaitu:

1. Bagaimana mengidentifikasi kecurigaan berita hoaks menggunakan algoritma K-means dan algoritma DBSCAN?
2. Bagaimana perbandingan hasil evaluasi antara *Sum of Squared Error* (SSE), *Silhouette Score*, dan *Davies-Bouldin Index* (DBI) dalam konteks membedakan kecurigaan berita hoaks dan non-hoaks?

1.3 Batasan Masalah

Untuk menjaga fokus pada tujuan penelitian, peneliti telah memutuskan untuk membatasi cakupan penelitian. Berikut adalah beberapa batasan yang telah ditetapkan:

1. Fokus pengaplikasian deteksi hoaks hanya di *platform* media sosial, yaitu twitter (X).
2. Data X yang digunakan untuk melakukan analisa adalah cuitan yang diunggah oleh seluruh pengguna di Indonesia dengan menggunakan kata kunci pilpres.
3. Rentang waktu data X yang diunggah dimulai dari 1 Januari hingga 21 Mei 2023.
4. Penelitian ini hanya memfokuskan identifikasi berita hoaks dengan membandingkan dua algoritma, yaitu algoritma DBSCAN dan algoritma K-Means.
5. Tahap evaluasi dari penelitian ini hanya mencakup metrik seperti *Sum of Squared Error* (SSE), *Sillhoutte Score*, dan *Davies-Bouldin Index* (DBI).
6. Fokus penelitian ini akan diarahkan pada perbandingan kinerja model algoritma DBSCAN dan K-Means dan metode evaluasi tanpa

mempertimbangkan aspek teknis implementasi model ke dalam sistem *deployment*.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Berdasarkan perumusan masalah yang telah ditetapkan, beberapa tujuan dari penelitian ini adalah sebagai berikut:

1. Mengidentifikasi kecurigaan berita hoaks menggunakan algoritma K-means dan algoritma dalam memprediksi kecurigaan berita hoaks.
2. Menentukan perbandingan hasil evaluasi antara *Sum of Squared Error* (SSE), *Silhouette Score*, dan *Davies-Bouldin Index* (DBI) dalam konteks membedakan kecurigaan berita hoaks dan non-hoaks.

1.4.2 Manfaat Penelitian

Penelitian yang dilakukan diharapkan mampu memberikan manfaat praktis dan teoritis, hal itu disampaikan sebagai berikut:

Manfaat Teoritis:

1. Memberikan pemahaman mengenai algoritma klustering, khususnya K-Means dan DBSCAN, serta kemampuan masing-masing algoritma dalam menangani deteksi kecurigaan berita hoaks.
2. Memungkinkan identifikasi tren hoaks yang muncul di platform X sehingga membantu dalam pemahaman evolusi dan penyebaran berita palsu.
3. Memberikan rekomendasi bagi para peneliti dan praktisi untuk memilih model *machine learning* yang paling sesuai dalam konteks deteksi hoaks pada berita.

Manfaat Praktis:

1. Memberikan pemahaman lebih dalam tentang risiko berita hoaks dan dampaknya terhadap kestabilan sosial sehingga dapat dilakukan penanganan yang efektif.

2. Memberikan kontribusi nyata dalam melindungi keamanan informasi serta memperkuat upaya dalam mengatasi hoaks di platform X yang terdapat di Indonesia.
3. Menjadi landasan untuk pengembangan kebijakan pemerintah atau langkah-langkah masyarakat dalam mengurangi penyebaran berita hoaks di Indonesia.

1.5 Sistematika Penulisan

1. BAB 1 PENDAHULUAN

Bab ini berisikan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

2. BAB II LANDASAN TEORI

Bab ini menjelaskan teori yang digunakan sesuai dengan topik dari penelitian yang akan dilakukan, serta membahas penelitian terdahulu yang menjadi sumber referensi untuk penelitian yang dilakukan saat ini.

3. BAB III METODOLOGI PENELITIAN

Bab ini membahas metode penelitian yang akan digunakan sebagai panduan dalam merancang tahap penelitian dari awal hingga akhir, seperti membuat alur proses yang akan dilakukan pada saat penelitian.

4. BAB IV ANALISIS DAN HASIL PENELITIAN

Bab ini menguraikan hasil yang didapatkan serta pembahasan yang lebih rinci sesuai dengan rumusan masalah yang telah diajukan.

5. BAB V SIMPULAN DAN SARAN

Bab ini berfokus pada kesimpulan utama dari penelitian yang telah dilakukan serta memberikan rekomendasi dan saran untuk penelitian mendatang.