

BAB I

PENDAHULUAN

A. Latar Belakang

Komunikasi merupakan proses seseorang atau kelompok orang untuk menciptakan suatu informasi maupun saling berbagi informasi baik verbal maupun non-verbal. Proses ini sangat penting bagi kehidupan manusia untuk menunjang berbagai aktivitas dan kebutuhannya. Sehingga dalam proses berkomunikasi terus mengalami perkembangan baik dari segi teknologi maupun dari diksi kata atau isi sebuah informasi yang bertambah. Teknologi sebagai media dalam berkomunikasi terus mengalami perubahan hingga pada sekarang ini telah ditemukan internet, kemampuan internet mendorong informasi menyebar dengan cepat dan menjangkau lebih banyak orang, sehingga informasi yang kita peroleh lebih *up to date*. Namun selain sisi positif teknologi ini, juga memiliki dampak negatif. Dimana dampak negatif yang dihasilkan informasi yang tidak seharusnya dipublikasikan atau milik privasi orang ataupun suatu berita tidak benar adanya menjadi konsumsi publik dan menimbulkan kegaduhan. Seperti contohnya kondisi-kondisi menjelang pemilihan presiden.

Twitter sebagai salah satu platform media sosial yang memanfaatkan perkembangan teknologi internet dan sekarang ini Twitter merupakan platform yang cukup banyak digunakan di berbagai kalangan. Menurut laporan dari *We are Social Meltwater* ada sekitar 60,2% menggunakan media sosial Twitter dari total penduduk Indonesia atau sekitar 167 juta pengguna (Meltwater, 2023). Pada media Twitter tak jarang pendapat yang diposting mengandung unsur negatif atau menjelekkan seseorang yang kemudian postingan tersebut dikonsumsi oleh publik.

Berdasarkan data di *website* Patroli Siber pada kurun waktu 3 tahun yaitu dari tahun 2020 hingga 2022 jumlah kasus pencemaran nama baik sekitar 6.401 kasus. Kasus-kasus yang sempat viral belakangan ini seperti kasus berikut; kasus pertama, sempat viral pencemaran nama baik Ikatan Dokter Indonesia (IDI) pada 16 Juli 2020 oleh terdakwa Jerinx pada narasi postingan di media sosial ia menyebut "Gara-gara bangga jadi kacung WHO, IDI, dan Rumah Sakit dengan seenaknya mewajibkan semua orang yang akan melahirkan tes Covid-19. Telah banyak bukti

jika hasil tes sering ngawur kenapa dipaksakan? Kalau hasil tes-nya bikin stres dan menyebabkan kematian pada bayi/ibunya, siapa yang tanggung jawab?”. Namun Jerinx berdalih hanya mengkritik semata dan mengajak diskusi, kemudian dirinya terancam hukuman 6 tahun penjara oleh putusan hakim. Selanjutnya pada kasus yang kedua, terjadi pada bulan Mei tahun 2022 yang mana artis ternama Nikita Mirzani melakukan tuduhan serta pencemaran nama baik kepada saudara Dito Mahendra. Isi unggahan posting narasinya sebagai berikut "Ini dua muka orang yang diduga melakukan penyekapan dan pemukulan secara sadis ke mantan supir bebegig sawah, yang dilakukan di rumah ibu kandangnya bebegig, kepada Kepolisian Indonesia harus adil dalam menangani kasus sadis ini". Nikita pun terancam hukuman 12 tahun penjara. Dan masih banyak kasus yang tengah viral sekarang ini.

Keyakinan untuk bebas berekspresi merupakan hak vital untuk diperjuangkan yang didasarkan pada pemahaman bersama bahwa bangsa dan rakyatnya dapat makmur dan berkembang jika terdapat forum yang bebas dan terbuka untuk berekspresi. Namun kebebasan berekspresi tunduk pada batasan-batasan tertentu sepanjang dilakukan tanpa melanggar hak orang lain. Pihak yang dirugikan karena dihina oleh orang lain berhak menuntut ganti rugi. Pencemaran nama baik menghasilkan *Character Assassination* yang dianggap pelanggaran terhadap hak asasi manusia (Muhammad Rizaldi, 2013). Mengetahui fakta tersebut, demi keberlangsungan berpendapat maupun kebebasan berekspresi demi memakmurkan rakyat dan memajukan bangsa tanpa menjatuhkan kehormatan seseorang. Indonesia pun memiliki undang-undang yang mengatur hubungan antar masyarakat didalam media elektronik dan pasal 27 ayat 3 undang-undang informasi dan transaksi elektronik (UU ITE) khusus menangani kasus pencemaran nama baik.

UU ITE dalam praktiknya terkadang menimbulkan banyak persepsi dan multitafsir, sehingga sering disalahgunakan oleh berbagai pihak. Menanggapi masalah yang timbul dari UU ITE, pemerintah telah merumuskan dan menerbitkan Surat Keputusan Bersama Menteri Komunikasi dan Informatika Republik Indonesia, Jaksa Agung Republik Indonesia dan Kepala Kepolisian Negara Republik Indonesia atau disingkat SKB 3 Menteri pada tahun 2021. Dari pedoman

tersebut dapat diketahui unsur penuduhan yang menjadi permasalahan dan perlunya pendalaman bahasa serta pemantauan dari suatu postingan. Pasal tersebut merupakan delik aduan yang mana kasus yang dilaporkan merupakan persepsi dari korban secara individu. Maka perlu juga mengidentifikasi secara general kata maupun kalimat yang dapat dikategorikan sebagai penuduhan maupun pencemaran.

Pada media sosial Twitter, identifikasi secara otomatis postingan Twitter secara komputasi termasuk kedalam cabang ilmu *Nature Language Processing* disingkat *NLP*. Telah banyak penelitian pada cabang *NLP*, seperti dilakukan oleh Petr Cerva (2014) berjudul *Investigation of latent semantic analysis for clustering of czech news articles*, tujuan dari penelitian ini adalah untuk menemukan cara yang tepat untuk mengelompokkan dokumen berbahasa Ceko yang sangat persuasif dan kaya akan kosa kata. Oleh karena itu, dalam penelitiannya menggunakan Latent Semantic Analysis (LSA), yang memungkinkan menemukan estimasi nilai rendah dari matriks istilah dokumen, yang menggambarkan frekuensi kemunculan suatu kata (*term*) dalam dokumen masukan. Dalam kasus bahasa infleksi, klasifikasi yang lebih rendah penting karena dua alasan. Pertama, ini memungkinkan untuk mengasosiasikan dimensi terkait dengan istilah yang memiliki arti yang sama: Matriks dokumen-term yang diperkirakan dianggap terlalu jarang dibandingkan dengan matriks dokumen-term "nyata", karena hanya berisi kata-kata yang benar-benar muncul di setiap dokumen, sementara (karena sinonim) mengandung kumpulan kata yang jauh lebih besar dari semua kata yang seharusnya. Kedua, reduksi dengan SVD diharapkan untuk menghilangkan noise dari matriks term-dokumen asli, yang berisi term semantik yang tidak berhubungan (*noisy*) (Rott & Cerva, 2014).

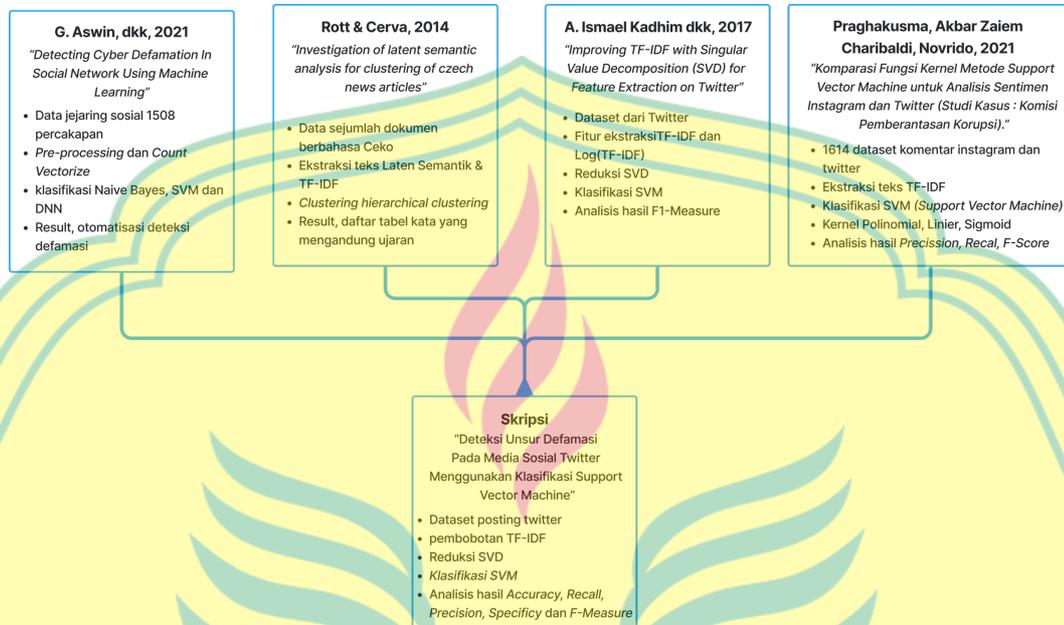
Dalam lingkup penelitian deteksi defamasi dilakukan oleh G. Aswin, dkk (2021), penelitian ini bertujuan mendeteksi secara otomatis postingan terkait defamasi dengan kategori *cyberbullying* di media sosial. Dengan jumlah dataset 1508 percakapan, 702 *cyberbullying*, 702 *non-cyberbullying*, 4654 jumlah kata berbeda, maksimal percakapan 663 karakter dan minimum percakapan 46 karakter. Sebelum dilakukan klasifikasi, data yang dikumpulkan akan di *pre-processing*, dan *Count Vectorize* saja. Pada proses klasifikasi dilakukan dengan dengan metode

Naive bayes, SVM dan *Deep Neural Network* (DNN) model. Dari penelitian ini menghasilkan kesimpulan respon lebih cepat dibandingkan deteksi *cyberbullying* secara manual (Pavitra et al., 2021).

Penelitian selanjutnya metode yang digunakan untuk meningkatkan hasil pengklasifikasian pada ekstraksi fitur teks, dilakukan oleh A. Ismael Kadhim dkk tahun 2017 dengan judul *Improving TF-IDF with Singular Value Decomposition (SVD) for Feature Extraction on Twitter*. Penelitian ini mengimplementasikan dua metode term frequency-inverse document frequency (TF-IDF) dan logaritma (TF-IDF) dengan teknik reduksi dimensi singular value decomposition (SVD). Makalah ini menyajikan metode baru yang menampilkan teknik preprocessing dan pengurangan dimensi yang efektif yang membantu ekstraksi fitur dengan menggunakan metode logaritma TF-IDF. Terakhir, hasil percobaan menunjukkan bahwa metode logaritma TF-IDF meningkatkan kinerja klasifikasi dokumen teks berbahasa Inggris. Hasil simulasi menunjukkan keunggulan dari algoritma yang diusulkan. Secara umum, TF-IDF dengan logaritma mengungguli TF-IDF tradisional sehubungan dengan metrik evaluasi. Pengukuran F1 meningkat dengan peningkatan jumlah fitur serta Log (TF-IDF) dengan SVD memiliki stabilitas yang lebih tinggi dengan peningkatan jumlah fitur (Alhaq et al., 2021).

Lalu dalam penelitian penggunaan klasifikasi SVM sebagai kategorisasi teks dapat diterapkan kernel untuk memisahkan data non-linier dan meningkatkan ke ruang dimensi yang lebih tinggi. Penelitian perbandingan setiap kernel dilakukan oleh Praghakusma dan Charibaldi pada tahun 2021, dengan judul *Komparasi Fungsi Kernel Metode Support Vector Machine untuk Analisis Sentimen Instagram dan Twitter (Studi Kasus : Komisi Pemberantasan Korupsi)*. Tujuan dari penelitian ini yaitu menerapkan dengan membandingkan kernel linier, kernel polinomial dan kernel sigmoid pada metode Support Vector Machine untuk klasifikasi analisis sentimen. Pada penelitian sebelumnya disebutkan disini RBF menghasilkan akurasi sebesar 81.32%, presisi sebesar 71.47%, dan recall sebesar 87.64. Kemudian hasil penelitian ini didapatkan bahwa kernel linier memiliki akurasi tertinggi sebesar 83.06%, presisi sebesar 91.04%, dan recall sebesar 89.70%, untuk kernel polinomial memiliki akurasi sebesar 81.45%, presisi sebesar 88.57%, dan recall

sebesar 91.17% sedangkan kernel sigmoid memiliki akurasi sebesar 79.83%, presisi sebesar 91.93%, dan recall sebesar 83.82%.



Gambar 1. 1 Pohon Penelitian

Berdasarkan penelitian yang telah dilakukan, penulis tertarik untuk melakukan penelitian yang menerapkan metode pembobotan TF-IDF dengan teknik reduksi *Singular Value Decomposition* sebagai ekstraksi fitur teks dengan klasifikasi SVM menggunakan kernel Linear pada identifikasi unsur defamasi di Twitter dengan judul: "Deteksi Unsur Defamasi Pada Media Sosial Twitter Menggunakan Klasifikasi *Support Vector Machine*". Adapun nanti data yang diambil merupakan nama 3 tokoh capres 2024 sebagai keyword ; Anies, Ganjar dan Prabowo. Proses pengambilan data di media sosial Twitter menggunakan teknik *Crawling* dengan librari Tweepy dan ditambah dataset dari database publik milik SafeNet.

B. Rumusan Masalah

Pada uraian di latar belakang, maka perumusan masalah di penelitian ini adalah :

1. Bagaimana membangun model klasifikasi untuk deteksi unsur defamasi pada media sosial Twitter menggunakan *Support Vector Machine*?

2. Bagaimana performa *Support Vector* dalam melakukan klasifikasi data defamasi media sosial Twitter dan data defamasi di SafeNet?
3. Bagaimana pengaruh penggunaan ekstraksi fitur teks TF-IDF dengan reduksi SVD pada model yang dibuat?

C. Batasan Masalah

Adapun batasan masalah yang ditentukan dari penelitian ini adalah sebagai berikut:

1. Data berupa postingan Twitter 2023 dan data dari SafeNet tentang kasus defamasi dari tahun 2018 hingga 2020 berbahasa Indonesia.
2. Pengumpulan data Twitter menggunakan Librari Tweepy dalam proses *Crawling data*.
3. Metode untuk ekstraksi fitur menggunakan TF-IDF dan teknik reduksi SVD.
4. Pada klasifikasi menggunakan *Support Vector Machine* dengan kernel Linear.
5. Bahasa pemrograman yang akan digunakan adalah *Python*.

D. Tujuan Penelitian

Berdasarkan rumusan masalah, maka tujuan penelitian ini adalah sebagai berikut :

1. Membangun model deteksi defamasi dengan melakukan klasifikasi SVM terhadap postingan Twitter dan data defamasi di SafeNet.
2. Mengetahui performa algoritma *Support Vector* dalam melakukan klasifikasi terhadap postingan Twitter dan data di SafeNet.
3. Mengetahui pengaruh penggunaan ekstraksi fitur TF-IDF dengan reduksi SVD pada model yang dibuat.

E. Manfaat Penelitian

Berdasarkan tujuan penelitian yang hendak dicapai, maka penelitian ini diharapkan mempunyai manfaat bagi personel kepolisian, masyarakat umum, maupun penulis. Adapun manfaat penelitian ini adalah sebagai berikut:

1. Bagi Personel Penegak Hukum

Memberikan informasi deteksi defamasi untuk pengawasan media sosial Twitter agar lebih cepat dan mudah.

2. Bagi Masyarakat umum

Membantu masyarakat dalam meningkatkan komunikasi yang lebih baik, menambah pengetahuan potensi pelanggaran mengenai pasal 27 ayat 3 UU ITE berdasarkan pedoman SKB 3 Menteri.

3. Bagi Penulis

Hasil perancangan Deteksi Defamasi dengan Metode Klasifikasi *Support Vector Machine* ini diharapkan dapat menambah wawasan dalam hal hukum, serta pengaruh kinerja dari ekstraksi teks TF-IDF dan teknik reduksi SVD dalam analisis defamasi.

