

BAB I PENDAHULUAN

Pada bab pendahuluan menjelaskan mengenai hal-hal yang menjadi latar belakang penelitian. Penelitian juga menjelaskan tentang identifikasi masalah, pembatasan masalah, dan rumusan masalah. Bab ini juga menjelaskan tentang tujuan dilakukannya penelitian dan manfaat penelitian. Berikut uraian dari latar belakang masalah, rumusan masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah dan sistematika penulisan yang dilakukan.

I.1 Latar Belakang

Pemilihan umum telah menjadi salah satu tonggak fundamental dalam sistem demokrasi, memberikan warga negara hak untuk memilih para pemimpin mereka. Dalam era digital yang semakin berkembang, media sosial telah mengambil peran sentral dalam memfasilitasi pertukaran informasi, pandangan, dan opini. Dalam konteks pemilihan umum, khususnya Pemilu 2024 di Indonesia, analisis sentimen yang terjadi melalui media sosial, terutama di platform Twitter, memiliki dampak yang sangat signifikan dalam memahami bagaimana masyarakat merespons berbagai aspek terkait pemilu tersebut. Sebelum pelaksanaan pemilu berlangsung, para calon presiden mempunyai strategi kampanye untuk memperbanyak pendukungnya, salah satunya adalah kampanye di media sosial dan *internet*.

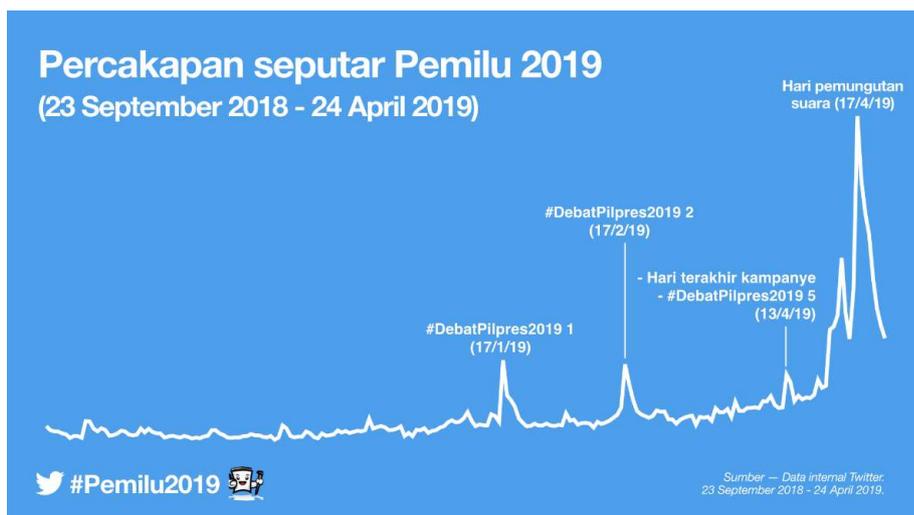
Strategi kampanye menggunakan media internet ini telah lama dilakukan dari tahun 2009 lalu, yang di mana beberapa calon presiden menyebarkan sosialisasi pemilu melalui *blog*. Calon presiden dan wakil presiden diusulkan oleh partai politik atau gabungan partai politik peserta pemilu. Dalam persiapan pemilu tahun 2024 nanti, beberapa partai politik telah mengusulkan nama-nama untuk calon presiden dan wakil presiden.

Twitter merupakan salah satu platform media sosial yang cukup populer untuk mengemukakan pendapat dan opini. Pengguna Twitter dapat mengirim dan membaca kicauan, atau "*tweet*" yang berbasis teks hingga 280 karakter. Sejak berdirinya Twitter, media sosial ini menjadi salah satu dari sepuluh situs yang paling sering dikunjungi di internet. Tercatat pada tahun 2022, pengguna Twitter

mencapai hampir 400 juta, dengan pengguna aktif setiap harinya mencapai 206 juta pengguna (Brian Dean, 2022).

Pemilihan platform media sosial yang tepat dalam analisis sentimen memiliki peran penting dalam memahami pandangan masyarakat terkait suatu topik. Dalam konteks Pemilu 2024 di Indonesia, pemilihan Twitter sebagai platform utama untuk analisis sentimen memiliki pertimbangan yang kuat. Alasan utama pemilihan Twitter adalah aksesibilitas data yang mudah diperoleh. Dengan memilih Twitter sebagai platform utama untuk analisis sentimen terkait Pemilu 2024, peneliti dapat memanfaatkan potensi data yang luas dan relevan untuk memperoleh wawasan mendalam tentang bagaimana masyarakat merespons, memahami sentimen yang beragam, serta merancang strategi kampanye yang lebih cerdas dan terarah mendekati proses pemilihan yang penting ini.

Data pada Gambar I.1 menunjukkan total aktivitas pengguna Twitter yang membahas topik pemilu 2019. Data ini bisa menjadi contoh bagaimana media sosial dapat mempengaruhi peran kampanye *online* pada pemilu 2024 dan dapat diprediksi bahwa akan ada banyak aktivitas lagi pada pemilu tahun 2024 nanti. Isu tentang calon Presiden potensial yang akan maju dalam pemilu 2024 mendatang sedang menjadi topik hangat yang sedang dibicarakan oleh publik saat ini, dan secara alami memunculkan berbagai reaksi dan pendapat.



Gambar I.1 Grafik Aktivitas pengguna Twitter seputar Pemilu 2019 (Agung Yudhawanata, 2019)

Menurut hasil penelitian yang dilakukan oleh Indikator Politik Indonesia, kepercayaan masyarakat terhadap sistem demokrasi menunjukkan tingkat yang sangat tinggi, mencapai 77,2%. Persentase ini mengalami peningkatan dibandingkan dengan survei sebelumnya yang membahas hal serupa. Penelitian ini melibatkan 1.200 responden yang diambil secara acak dari seluruh Provinsi, dengan distribusi yang proporsional. Tingginya tingkat antusiasme dari para responden menunjukkan minat besar masyarakat untuk terlibat dalam demokrasi. Dengan menganalisis sentimen, strategi kampanye calon presiden dapat dirancang secara efektif berdasarkan dukungan publik yang diukur.

Banyaknya informasi *tweet* tiap harinya membuat manusia kesulitan dalam mengolahnya. Selain itu juga, belum adanya otomatisasi penilaian opini pada platform Twitter membuat pengolahan informasi menjadi sulit. Hal inilah menjadikan sulitnya menilai sentimen dan opini masyarakat Indonesia yang memakai media sosial secara manual, sehingga perlu diterapkannya otomatisasi yang dapat memudahkan hal tersebut (Farisi & Muslim, 2020). Salah satu metode yang dapat digunakan untuk memahami opini pada Twitter yaitu analisis sentimen.

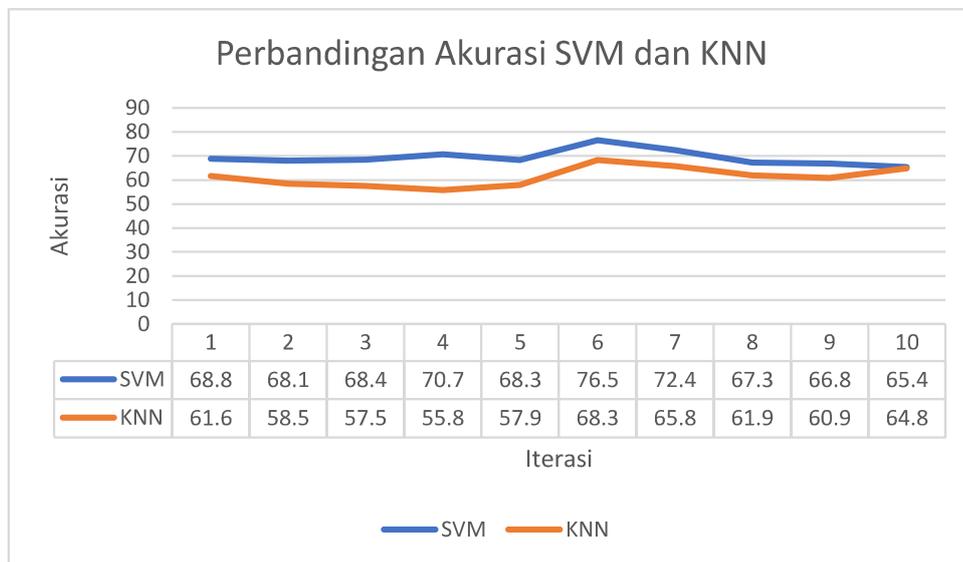
Analisis sentimen telah menjadi sebuah instrumen penting untuk mengukur dan memahami opini serta pandangan masyarakat terkait suatu topik tertentu. Dalam hal Pemilu 2024, analisis sentimen memiliki potensi besar untuk memberikan wawasan mendalam tentang cara masyarakat merespons calon kandidat, partai politik, dan isu-isu yang menjadi sorotan dalam proses pemilihan.

Analisis sentimen atau yang sekarang sering kali disebut dengan *opinion analysis* adalah salah satu jenis *Natural Language Processing* (NLP) mengenai proses pemahaman suka dan tidak suka yang diolah dari kalimat opini ataupun komentar menjadi sebuah informasi sentimental yang akan dikategorikan menjadi negatif, positif, atau netral (Budiharto & Meiliana, 2018; Naf'an dkk., 2019). Analisis sentimen pada pemilihan umum presiden dan wakil presiden ini digunakan untuk menganalisis dan memprediksi hasil pemilihan umum berdasarkan opini dan komentar masyarakat mengenai pemilu di platform Twitter.

Tantangan utama dalam menghadapi analisis sentimen pada skala besar adalah volume data yang melimpah, terutama di platform media sosial yang

menghasilkan ratusan ribu hingga jutaan pesan setiap harinya. Analisis manual menjadi tidak mungkin dan tidak efisien dalam konteks ini. Dalam hal ini, algoritma komputasional, khususnya K-Nearest Neighbors (KNN), memiliki potensi untuk membantu mengatasi tantangan ini dengan pendekatan klasifikasi otomatis.

Pemodelan KNN adalah pemodelan klasifikasi paling dasar dan sederhana yang membutuhkan sedikit atau bahkan tidak ada sama sekali pengetahuan sebelumnya tentang sifat-sifat pada data. KNN dipandang sebagai pemodelan yang memakan proses komputasi dan alokasi memori yang banyak karena pemakaiannya (Yu & Nwet, 2020). Pada penelitian yang telah dilakukan oleh Firmansyah (2020) pada jurnal yang berjudul “*Comparing Sentiment Analysis of Indonesian Presidential Election 2019 with Support Vector Machine and K-Nearest Neighbor Algorithm*”, pemodelan KNN dibandingkan dengan pemodelan *Support Vector System* (SVM) untuk melihat seberapa besar akurasi dari kedua pemodelan tersebut. *Data* yang dipakai adalah data pemilu Indonesia di tahun 2019 pada Twitter, dan diperoleh hasil yang ditampilkan pada Gambar I.2:



Gambar I.2 Perbandingan Akurasi SVM dan KNN (Firmansyah dkk., 2020)

Pada Gambar I.2 penelitian dilakukan dengan menguji *dataset* Twitter sebanyak 10 kali. Hasil penelitian oleh Firmansyah disimpulkan bahwa KNN memiliki rata-rata akurasi yang rendah dibandingkan dengan pemodelan SVM, dengan nilai maksimal dari SVM sebesar 76.5% dan nilai maksimal dari KNN sebesar 65.8%.

Pada penelitian yang serupa (Yu & Nwet, 2020), pemodelan KNN menghasilkan tingkat akurasi yang rendah dengan pemodelan lainnya, hal ini menyebabkan pemodelan KNN disebut sebagai pemodelan yang kurang efektif.

Beberapa penelitian menggabungkan model KNN dengan model lainnya untuk menghasilkan peningkatan akurasi. Pada penelitian yang dilakukan oleh (Giovani dkk., 2020), mereka membuat model kombinasi dari KNN dan *Particle Swarm Optimization* (PSO) dengan hasil awal tanpa bantuan PSO sebesar 70.42% dan hasil dengan bantuan PSO sebesar 77.21%. Penelitian lainnya yang dilakukan oleh (Istia & Purnomo, 2018) juga mengukur tingkat akurasi KNN yang menggabungkan penggunaan model PSO yang ditampilkan pada Tabel I.1.

Tabel I.1 Perbandingan KNN dengan KNN-PSO

<i>K-Value</i>	KNN		KNN-PSO	
	Akurasi	AUC	Akurasi	AUC
2	90.02%	0.973	92.67%	0.975
3	91.04%	0.960	93.08%	0.948
4	89.82%	0.940	90.63%	0.943
5	90.22%	0.920	91.45%	0.923
6	89.41%	0.911	90.02%	0.911
7	90.22%	0.917	90.84%	0.898
8	89.21%	0.911	90.22%	0.898
9	90.63%	0.903	90.63%	0.903
10	89.41%	0.895	90.43%	0.891

Tabel I.1 menjelaskan tentang pengujian data dengan *K-Value* sebanyak 10 kali. Hasil pengujian pada KNN mendapatkan hasil tertinggi pada angka 91.04%, dan tingkat akurasi pada model yang menggabungkan KNN-PSO mendapat hasil 93.08%. Dapat disimpulkan bahwa pemodelan KNN-PSO mendapat hasil akurasi yang lebih besar daripada KNN.

Fokus utama dari penelitian ini adalah untuk membuktikan performa K-Nearest Neighbors (KNN) dengan integrasi Particle Swarm Optimization (PSO) dalam konteks analisis sentimen terkait Pemilu 2024 di Twitter. Analisis sentimen telah menjadi sebuah alat penting untuk menggali pandangan dan opini masyarakat terhadap isu-isu yang berkaitan dengan pemilihan umum. Dalam kerangka ini, tujuan utama penelitian ini adalah untuk memperdalam pemahaman tentang

kemampuan dan keefektifan metode KNN ketika ditingkatkan dengan PSO dalam mengklasifikasikan opini masyarakat pada dataset Twitter terkait Pemilu 2024.

Kemampuan KNN dalam mengklasifikasikan data berdasarkan keterdekatannya dengan data latihan telah terbukti berguna dalam berbagai aplikasi analisis sentimen. Namun, performa KNN dapat dipengaruhi oleh faktor-faktor tertentu seperti pemilihan jumlah tetangga terdekat yang optimal. Di sinilah PSO datang sebagai pendekatan yang menarik. PSO, dengan sifat optimisasi berdasarkan gerakan kelompok, dapat membantu mengoptimalkan pemilihan parameter-parameter yang mempengaruhi performa KNN dalam mengklasifikasikan sentimen.

I.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka dapat dirumuskan beberapa permasalahan yang ada, yaitu:

1. Apakah algoritma K-NN dikombinasikan dengan optimisasi PSO akan menghasilkan performa yang baik?

I.3 Tujuan Penelitian

Adapun tujuan dari hasil perumusan masalah yang dibuat ialah sebagai berikut:

1. Mengevaluasi apakah algoritma K-NN yang dikombinasikan dengan optimisasi PSO menghasilkan performa yang lebih baik.

I.4 Batasan Penelitian

Agar permasalahan dapat mencapai tujuan yang telah ditentukan, maka dibuat batasan-batasan sebagai berikut:

1. Analisis sentimen yang digunakan merupakan hasil dari metode KNN yang dikombinasikan dengan PSO.
2. Data yang diambil merupakan 3 kandidat terkuat pada pemilu 2024 saja.
3. *Dataset* yang diambil merupakan data *tweet* dari rentang tanggal 1 Januari 2023 sampai 14 Maret 2023

I.5 Manfaat Penelitian

Beberapa manfaat dari penelitian ini adalah sebagai berikut:

1. Manfaat untuk peneliti, menambah wawasan dalam proses analisis sentimen menggunakan kombinasi model *K-Nearest Neighbor* dan *Particle Swarm Optimization (PSO)*.
2. Manfaat bagi pemerintah, diharapkan dapat memberikan informasi terkait opini masyarakat tentang calon-calon presiden yang diselenggarakan sebelum pemilu untuk dijadikan strategi kampanye.
3. Manfaat bagi akademik, diharapkan dapat memberikan kontribusi dan referensi untuk penelitian lebih lanjut.

I.6 Sistematika Penulisan

Tugas akhir ini diuraikan dengan sistematika penulisan sebagai berikut:

Bab 1 Pendahuluan

Bab ini berisi uraian atau penjelasan mengapa penelitian ini dilakukan yang meliputi latar belakang, perumusan masalah, tujuan penelitian, batasan penelitian, serta manfaat yang diperoleh pada penelitian.

Bab II Tinjauan Pustaka

Bab ini berisi penjelasan terkait hasil studi pustaka penelitian terdahulu, dasar teori yang berkaitan dengan permasalahan yang diambil. Lalu penjelasan mengenai tahap persiapan, algoritma beserta mode lain yang terkait.

Bab III Metodologi Penelitian

Bab ini menjelaskan mengenai metode langkah-langkah dalam penelitian, dari tahap pengumpulan data, pengolahan data, algoritma pengolahan data sampai dengan tahap akhir dari penelitian

Bab IV Analisis dan Perancangan

Bab ini menjelaskan mengenai hasil pengerjaan dan pembahasan dari studi kasus yang sudah dipaparkan sebelumnya. Dalam bab ini

penelitian melakukan analisis *preprocessing* dan menjelaskan metode *data mining* yang dipakai dalam penelitian.

Bab V Hasil dan Evaluasi

Bab ini menjelaskan mengenai hasil yang diperoleh selama perancangan dan disertai dengan analisis.

Bab VI Kesimpulan dan Saran

Bab ini menyimpulkan dari seluruh analisis tugas akhir yang dilakukan serta saran yang dapat penelitian