BAB I

PENDAHULUAN

1.1. Latar Belakang Penelitian

Program Koperasi Desa/Kelurahan Merah Putih (KDKMP) merupakan upaya pemerintah dalam memperkuat kelembagaan ekonomi di tingkat desa dan kelurahan yang memiliki peran strategis dalam aktivitas ekonomi dan pemberdayaan masyarakat (Noho, 2025). Koperasi dalam konteks tersebut berperan dalam mendorong pemerataan ekonomi dan kemandirian masyarakat di tingkat lokal (Irawan, 2021). Namun, keberhasilan KDKMP tidak hanya ditentukan oleh terbentuknya koperasi, tetapi juga oleh kemampuan kelembagaan dalam menjalankan fungsi usaha dan melibatkan anggota secara berkelanjutan (Nugraha, 2026). Kondisi tersebut menjadi perhatian dalam percakapan publik di media sosial *X* yang memunculkan beragam pandangan mengenai pelaksanaan program KDKMP. Keberagaman opini dalam jumlah data yang besar memerlukan pendekatan klasifikasi teks untuk mengidentifikasi kecenderungan sentimen secara sistematis.

Pemerintah meluncurkan Program Koperasi Desa/Kelurahan Merah Putih (KDKMP) sebagai upaya memperkuat kelembagaan ekonomi desa dan kelurahan. Program ini didukung regulasi utama, yaitu Petunjuk Pelaksanaan Menteri Koperasi Nomor 1 Tahun 2025 tentang Pembentukan Koperasi Desa/Kelurahan Merah Putih, Instruksi Presiden Nomor 9 Tahun 2025 tentang Percepatan

Pembentukan 80.000 Koperasi Desa/Kelurahan Merah Putih, serta Peraturan Menteri Hukum Nomor 13 Tahun 2025 tentang Pengesahan Koperasi (Republik Indonesia, 2025b; Kementerian Koperasi Republik Indonesia, 2025a; Kementerian Hukum dan Hak Asasi Manusia Republik Indonesia, 2025). Dalam pelaksanaannya, pembentukan KDKMP dilakukan melalui tiga mekanisme, yaitu pendirian koperasi baru, pengembangan koperasi yang telah ada, dan revitalisasi koperasi yang tidak aktif. Ketiga mekanisme tersebut merupakan pilihan dalam proses pembentukan KDKMP dan tidak menunjukkan jenis badan hukum yang berbeda maupun menjadi bukti bahwa koperasi telah menjalankan kegiatan operasionalnya (Kementerian Koperasi Republik Indonesia, 2025a). Seluruh mekanisme tersebut diarahkan untuk menerapkan jati diri koperasi (*cooperative identity*) sebagaimana dirumuskan *International Cooperative Alliance (ICA)*, meliputi keanggotaan sukarela, partisipasi ekonomi anggota, pendidikan, kerja sama antarkoperasi, dan kepedulian terhadap masyarakat (Kementerian Koperasi Republik Indonesia, 2025a).

Keberhasilan Program KDKMP tidak hanya berkaitan dengan jumlah koperasi yang berhasil dibentuk, tetapi juga dengan kemampuanya menciptakan nilai ekonomi, meningkatkan kesejahteraan, serta menjadi motor penggerak pembangunan ekonomi berkelanjutan (Undang-Undang Nomor 25 Tahun 1992). Pembentukan KDKMP yang banyak didorong melalui kebijakan pemerintah turut menimbulkan perdebatan mengenai kesesuaiannya dengan jati diri koperasi. Perdebatan tersebut berkaitan dengan penerapan prinsip

keanggotaan sukarela, pengendalian secara demokratis, partisipasi ekonomi anggota, serta tata kelola koperasi yang menempatkan anggota sebagai pemilik sekaligus pengguna koperasi (Kementerian Koperasi Republik Indonesia, 2025a). Namun demikian, pendekatan *top-down* dalam pembentukan KDKMP tidak dapat secara langsung diartikan sebagai bukti bahwa seluruh koperasi yang terbentuk telah atau tidak menerapkan prinsip koperasi. Penilaian terhadap penerapan jati diri koperasi memerlukan pengamatan terhadap praktik kelembagaan dan operasional masing-masing koperasi. Selain hal tersebut, keberlanjutan KDKMP juga berkaitan dengan penguatan kapasitas kelembagaan dan pengembangan usaha agar koperasi mampu menjalankan fungsi ekonominya secara berkelanjutan (Nugraha, 2026). Kondisi tersebut menunjukkan adanya beragam persepsi masyarakat terhadap pelaksanaan Program KDKMP yang dapat dikaji melalui klasifikasi sentimen.

Pelaksanaan Program KDKMP didukung oleh pendanaan yang berasal dari beberapa sumber. Berdasarkan Instruksi Presiden Nomor 9 Tahun 2025 yang ditetapkan pada 27 Maret 2025, pendanaan dalam rangka percepatan pembentukan 80.000 Koperasi Desa/Kelurahan Merah Putih dapat bersumber dari APBN, APBD, APBDes, serta didukung likuiditas dari Himpunan Bank Milik Negara atau Himbara (Republik Indonesia, 2025b). Selain itu, pemerintah menyediakan skema pinjaman bagi KDKMP melalui Peraturan Menteri Keuangan Nomor 49 Tahun 2025. Dalam skema tersebut, pinjaman yang dapat diberikan kepada setiap koperasi paling banyak sebesar Rp3 miliar, dengan tingkat bunga 6 persen per tahun dan jangka waktu paling lama 72 bulan

(Kementerian Keuangan Republik Indonesia, 2025a). Ketentuan tersebut menunjukkan batas maksimal pembiayaan yang dapat diberikan melalui skema pinjaman, sedangkan pemanfaatannya tetap mengikuti ketentuan dan mekanisme yang berlaku. Dengan demikian, besaran plafon pinjaman tersebut tidak menunjukkan bahwa seluruh koperasi yang menjadi target program memperoleh pembiayaan dalam jumlah yang sama. Peraturan Menteri Keuangan Nomor 49 Tahun 2025 kemudian dicabut dan digantikan oleh Peraturan Menteri Keuangan Nomor 15 Tahun 2026 yang berlaku sejal 1 April 2026 (Kementerian Keuangan Republik Indonesia, 2026).

Gambar 1.1 menyajikan gambaran program berdasarkan materi sosialisasi Kementerian BUMN tanggal 16 April 2025. Materi tersebut memuat target 80.000 koperasi, modal awal sekitar Rp5 miliar per koperasi, serta perkiraan kebutuhan pendanaan sekitar Rp400 triliun (Kementerian BUMN, 2025). Angka tersebut merupakan kebutuhan pendanaan yang disajikan dalam materi program dan bukan realisasi anggaran.



Gambar 1.1 Overview Program Koperasi Desa/Kelurahan Merah Putih

Sumber: Kementerian BUMN (2025), materi sosialisasi Inpres Nomor 9 Tahun 2025, 16 April 2025.

Skala dan dukungan tersebut menjadikan Program Koperasi Desa/Kelurahan Merah Putih sebagai kebijakan yang memperoleh perhatian luas dan menjadi perbincangan publik.

Ruang digital ini terus berkembang seiring meningkatnya penetrasi internet di Indonesia, yang mencapai 80,66 persen atau sekitar 229,4 juta pengguna dengan media sosial sebagai aktivitas daring yang paling banyak dilakukan (APJII, 2025). Kondisi tersebut menjadikan media sosial sumber data yang relevan untuk mengidentifikasi persepsi publik terhadap kebijakan, termasuk Program KDKMP (Dahana, 2024). *X* menjadi salah satu platform yang banyak digunakan sebagai ruang diskusi publik dan memungkinkan pengguna menyampaikan opini secara cepat (Murthy, 2024; Kumar & Singh, 2022). *Platform* ini menyediakan data teks yang lebih terstruktur dibandingkan media sosial lain yang didominasi konten visual, didukung *fitur advanced search* dan *Tweet Harvest* untuk pengumpulan data tanpa akses *API* berbayar (Kumar & Singh, 2022; Satria, 2023), dengan estimasi jangkauan pengguna di Indonesia mencapai 22,9 juta (Kemp, 2025). Pertimbangan tersebut mendasari pemilihan *X* sebagai sumber data penelitian ini, dengan data dikumpulkan melalui *web scraping* menggunakan kata kunci terkait Program KDKMP pada rentang 30 April hingga 31 Mei 2026. Proses tersebut berhasil menjaring 8.954 unggahan dalam sebulan, menunjukkan bahwa topik ini cukup ramai diperbincangkan pengguna *X*.

Data opini dari *X* memiliki karakteristik tidak terstruktur (*unstructured text*) yang mengandung *hashtag*, *mention, URL*, emoji, singkatan, dan bahasa

informal (Das *et al.*, 2024). Karakteristik tersebut berpotensi menimbulkan *noise*, sehingga data perlu melalui tahapan *preprocessing*, meliputi *case folding*, *cleaning text*, *tokenizing*, normalisasi, *stopword removal*, dan *stemming* menggunakan pustaka *Sastrawi*, untuk meningkatkan kualitas representasi fitur (Purbo, 2019; Hickman *et al.*, 2022). Pada tahap *cleaning*, *URL*, *mention*, tanda baca, dan karakter khusus dihapus, sedangkan simbol *hashtag (#)* dihapus, sedangkan kata yang menyertainya dipertahankan, dan *emoji* dihapus. Selanjutnya, normalisasi dilakukan untuk mengubah *slang word*, singkatan, dan variasi penulisan menjadi bentuk baku sebelum proses *stemming* menggunakan pustaka *Sastrawi*, sehingga kata berimbuhan dapat direduksi menjadi kata dasar secara lebih akurat. Namun demikian, karena *stemming* berpotensi menghilangkan sebagian konteks kata, penelitian ini membandingkan skenario dengan dan tanpa *stemming* untuk menguji pengaruhnya terhadap performa klasifikasi sentimen (Hickman *et al.*, 2022).

Opini yang telah melalui tahap *preprocessing* , dianalisis ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral (Purnamasari, 2023). Ketiga kategori tersebut dapat ditemukan secara nyata dalam percakapan pengguna *X*, sebagaimana disajikan pada Tabel 1.1 berikut.

Tabel 1. 1 Contoh Percakapan Pengguna Media Sosial *X*

Kategori	Contoh Percakapan
Positif	Keren abis Koperasi Desa Merah Putih, bantu rakyat dapat kerja nyata.
Negatif	Koperasi Merah Putih sama sekali tidak memenuhi kriteria koperasi.
Netral	Menteri Keuangan memberikan tanggapan mengenai pengalokasian dana APBN untuk membiayai pegawai, termasuk Koperasi Desa.

Keberagaman sentimen tersebut, disertai volume unggahan yang besar, menyebabkan identifikasi sentimen secara manual menjadi kurang efisien dan rentan bias. Kondisi ini membutuhkan metode yang mampu mengonversi teks menjadi representasi numerik agar dapat diproses secara otomatis oleh algoritma klasifikasi. Penelitian ini menggunakan *Term Frequency–Inverse Document Frequency (TF-IDF)* (Manning *et al.*, 2008) dengan pendekatan *unigram* dan *bigram*. Pendekatan *unigram* merepresentasikan setiap kata sebagai fitur terpisah, sedangkan *bigram* mempertahankan pasangan kata berurutan sehingga berpotensi menangkap konteks frasa, terutama pada negasi seperti "tidak efektif", meskipun konsekuensinya jumlah fitur menjadi lebih besar (Indarbensyah & Rochmawati, 2021). Performa kedua representasi ini sangat bergantung pada karakteristik data, sehingga penelitian ini membandingkan keduanya secara empiris. Representasi fitur tersebut selanjutnya diklasifikasikan menggunakan algoritma *Multinomial Naïve Bayes*, metode berbasis probabilitas yang sederhana, cepat, dan stabil untuk data diskrit berbasis frekuensi kata (Fitrianti & Yudhistira, 2025; Mahendra *et al.*, 2025). Algoritma ini terbukti efektif pada berbagai penelitian analisis sentimen sebelumnya (Safira & Hasan, 2023; Verawati & Jaelani, 2024; Wati *et al.*, 2023).

Penelitian terdahulu menunjukkan variasi performa antar algoritma pada topik serupa. Fauzi, Andini, dan Muslihudin (2025) serta Fauziah *et al.* (2025) mendapati *Support Vector Machine (SVM)* unggul pada beberapa skenario meski terbatas dalam menangani bahasa informal, sedangkan *Naïve*

Bayes tampil lebih stabil dan efisien. Qodriyah *et al.* (2026) menemukan perbedaan karakteristik sentimen antara platform *X* dan *TikTok* menggunakan *TF-IDF* dan *SVM*. Dermawan (2026) memperoleh akurasi 90 persen menggunakan *TF-IDF*, pelabelan berbasis *leksikon,* dan *Random Forest*, tetapi nilai *Macro F1-score* yang hanya 0,63 menunjukkan kemampuan model belum seimbang dalam mengenali setiap kelas sentimen.

Penelitian-penelitian tersebut umumnya masih berfokus pada algoritma *Support Vector Machine* atau *Random Forest*, menempatkan *Multinomial Naïve Bayes* sekadar sebagai metode pembanding. Evaluasi performa pada penelitian terdahulu juga lebih menitikberatkan nilai akurasi, belum banyak menggunakan metrik yang lebih komprehensif seperti *Macro F1-score*, *Weighted F1-score, Balanced Accuracy,* dan *Confusion Matrix*. Perbandingan pengaruh *stemming* serta representasi *TF-IDF unigram* dan *bigram* terhadap performa *Multinomial Naïve Bayes* pada opini publik mengenai Program KDKMP juga masih terbatas jumlahnya. Kesenjangan tersebut mendasari penelitian ini mengangkat tema **"Klasifikasi Sentimen Opini Publik terhadap Program Koperasi Desa/Kelurahan Merah Putih pada Media Sosial *X* Menggunakan *Multinomial Naïve Bayes*",** dengan menerapkan *metodologi Cross-Industry Standard Process for Data mining (CRISP-DM)* dan membandingkan empat skenario klasifikasi berdasarkan kombinasi *TF-IDF unigram* atau *bigram* serta ada dan tidaknya *stemming*. Hasil klasifikasi yang diperoleh diharapkan mampu menggambarkan kecenderungan persepsi

masyarakat, sekaligus menjadi masukan bagi pemerintah dan pemangku kepentingan dalam mengevaluasi pelaksanaan program.

IKOPIN University

1.2. Identifikasi Masalah

Berdasarkan latar belakang yang telah diuraikan, maka identifikasi masalah dalam penelitian ini dirumuskan sebagai berikut:

1. Bagaimana distribusi sentimen positif, negatif, dan netral serta karakteristik kata-kata dominan dalam opini penggunaan media sosial *X* mengenai Program Koperasi Desa/Kelurahan Merah Putih
2. Bagaimana penerapan *stemming* dan tanpa *stemming* serta penggunaan representasi fitur *TF-IDF* berbasis *unigram* dan *bigram* terhadap performa klasifikasi sentimen menggunakan algoritma *Multinomial Naïve Bayes*
3. Skenario manakah yang mengshasilkan performa klasifikasi sentimen terbaik sebagai nilai *Macro F1-score* sebagai metrik utama, *serta accuracy, precision, recall, Weighted F1-score*, *Balanced Accuracy*, dan *Confusion Matrix* sebagai metrik pendukung
4. Bagaimana pola dan penyebab kesalahan klasifikasi model terbaik melalui *Confusion Matrix* guna mendokumentasikan keterbatasan performa model

IKOPIN University

1.3. Maksud dan Tujuan Penelitian

1.3.1. Maksud Penelitian

Maksud dari penelitian ini adalah melakukan klasifikasi sentimen opini publik terhadap Program Koperasi Desa/Kelurahan Merah Putih pada media sosial *X* menggunakan pembobotan *Term Frequency–Inverse Document Frequency (TF-IDF)* dan algoritma *Multinomial Naïve Bayes* . Penelitian ini bertujuan mengolah data opini yang tidak terstruktur menjadi informasi yang sistematis melalui tahapan pengumpulan data, *preprocessing*, representasi fitur, pemodelan, dan evaluasi sehingga dapat menggambarkan kecenderungan persepsi masyarakat terhadap Program Koperasi Desa/Kelurahan Merah Putih.

1.3.2. Tujuan Penelitian

Berdasarkan identifikasi masalah diatas maka yang menjadi tujuan dari penelitian ini adalah sebagai berikut:

1) Menganalisis distribusi sentimen positif, negatif, dan netral serta karakteristik kata-kata dominan dalam opini pengguna media sosial *X* mengenai Program Koperasi Desa/Kelurahan Merah Putih.
2) Menganalisis penerapan *stemming* dan tanpa *stemming* serta penggunaan representasi fitur *TF-IDF* berbasis *unigram* dan *bigram* terhadap performa klasifikasi sentimen menggunakan algoritma *Multinomial Naïve Bayes*.
3) Menentukan skenario klasifikasi sentimen terbaik berdasarkan nilai *Macro F1-score* sebagai metrik utama, serta *accuracy*, *precision*,

recall, *Weighted F1-score*, *Balanced Accuracy*, dan *Confusion Matrix* sebagai metrik pendukung.

4) Mengetahui pola dan penyebab kesalahan klasifikasi model terbaik melalui *Confusion Matrix* guna mendokumentasikan keterbatasan performa model

1.4. Kegunaan Penelitian

1.4.1. Kegunaan Teoritis

a. Bagi Peneliti, diharapkan dapat memberikan kontribusi empiris bagi pengembangan kajian sains data, khususnya di bidang *text mining*, natural language processing, dan analisis sentimen berbahasa Indonesia melalui perbandingan empat skenario klasifikasi *TF-IDF* terhadap performa algoritma *Multinomial Naïve Bayes* , sehingga memberikan gambaran pengaruh variasi *preprocessing* dan representasi fitur terhadap hasil klasifikasi.

b. Bagi Peneliti Selanjutnya, diharapkan dapat menjadi salah satu referensi empiris bagi penelitian selanjutnya yang membahas klasifikasi sentimen pada media sosial, terutama pada data teks berbahasa Indonesia. Hasil penelitian dapat digunakan sebagai bahan perbandingan untuk mengembangkan penelitian dengan algoritma lain, pendekatan *preprocessing* yang berbeda, periode pengambilan data yang lebih luas, atau sumber data dari platform digital lainnya

IKOPIN University

1.4.2. Kegunaan Praktis

a. Bagi Pemerintah, diharapkan dapat memberikan informasi pendukung mengenai kecenderungan sentimen positif, negatif, dan netral dalam percakapan pengguna media sosial *X* mengenai Program Koperasi Desa/Kelurahan Merah Putih. Informasi mengenai distribusi sentimen, kata-kata dominan, serta pola opini yang teridentifikasi dapat digunakan sebagai masukan awal untuk memahami isu yang berkembang dalam ruang publik digital. Namun, hasil penelitian tidak dimaksudkan sebagai satu-satunya dasar dalam mengevaluasi kebijakan atau mengambil keputusan karena data media sosial *X* tidak mewakili seluruh masyarakat Indonesia.
b. Bagi Pengelola dan pelaksana Program Koperasi Desa/Kelurahan Merah Putih, diharapkan dapat memberikan gambaran mengenai topik yang memperoleh perhatian, dukungan, pertanyaan, atau kritik dari pengguna media sosial *X*. Informasi tersebut dapat dimanfaatkan sebagai bahan pendukung dalam menyusun komunikasi publik, meningkatkan penyampaian informasi mengenai program, serta mengenali isu yang memerlukan penjelasan lebih lanjut kepada masyarakat.
c. Bagi Masyarakat, diharapkan penelitian ini dapat memberikan gambaran mengenai kecenderungan opini publik terhadap Program Koperasi Desa/Kelurahan Merah Putih , sehingga dapat

meningkatkan pemahaman masyarakat terhadap kebijakan yang sedang dijalankan oleh pemerintah.

IKOPIN University