

BAB I PENDAHULUAN

1.1 Latar Belakang

Pada era perkembangan teknologi yang semakin canggih saat ini, gaya hidup manusia diiringi dengan media yang serba digital. Perkembangan teknologi digital ini membuat masyarakat mulai beralih dari media lama menjadi media digital. Adanya media digital tersebut telah mengubah cara berjalannya arus informasi di seluruh dunia, termasuk di Indonesia. Kemunculan internet dan berbagai jenis media digital seperti media sosial telah memfasilitasi penyebaran akses informasi dengan cakupan lebih luas dan cepat kepada masyarakat. Hal ini memungkinkan penyebaran opini publik secara masif dan lebih terbuka [1].

Media sosial adalah sebuah sarana daring yang dapat memfasilitasi interaksi sosial dan pertukaran konten antar penggunanya [2]. Indonesia merupakan salah satu dari sekian banyak negara yang penduduknya mayoritas telah menggunakan media sosial. Pada zaman saat ini, media sosial telah menyatu dengan gaya hidup masyarakat Indonesia. Rata-rata waktu penggunaan sosial media masyarakat Indonesia lebih dari 3 jam per harinya [3]. Sosial media juga menjadi salah satu panggung besar dalam diskusi dan debat mengenai kebijakan-kebijakan pemerintah. Beberapa *platform* seperti YouTube, X, Instagram, dan Facebook tidak hanya menjadi tempat berbagi informasi, tetapi juga menjadi tempat untuk mengekspresikan opini publik terhadap kondisi negara Indonesia saat ini. Kebijakan-kebijakan yang dikeluarkan pemerintah selalu menjadi perbincangan hangat oleh masyarakat. Hal tersebut mengindikasikan beragam perspektif masyarakat terhadap isu-isu sosial, ekonomi, dan kultural yang berkembang.

Berbagai video yang membahas kebijakan pemerintah sering kali menarik perhatian besar dari masyarakat. Setiap kebijakan yang dikeluarkan pemerintah pasti menjadi perhatian masyarakat, terlebih lagi setelah proses demokrasi seperti Pemilu. Misalnya seperti kebijakan kenaikan Uang Kuliah Tunggal (UKT) dan Tabungan Perumahan Rakyat, kerap memicu beragam reaksi publik yang mencerminkan perasaan dan opini mereka tentang dampak dari keputusan tersebut. Kenaikan UKT dianggap memberatkan mahasiswa dan keluarga berpenghasilan

rendah, sementara kebijakan tabungan perumahan rakyat sering diperdebatkan terkait efektivitasnya dalam memenuhi kebutuhan hunian yang layak. Setelah pengumuman kebijakan baru, masyarakat cenderung merespons unggahan video tersebut dengan menyampaikan perasaan dan opini mereka tentang dampak adanya kebijakan tersebut, memberikan kritik, serta berbagai pandangan lainnya. Interaksi yang beragam ini mencerminkan dinamika sosial yang terus berkembang dalam proses demokrasi, di mana opini masyarakat menjadi indikator kritis untuk memahami fluktuasi opini publik dan implikasinya terhadap proses demokrasi di Indonesia. Ada yang mengutarakan perasaan senangnya terkait bentuk kebijakan yang akan ditetapkan pemerintah, dan ada juga yang menyampaikan perasaan marahnya, serta tak sedikit juga yang mengutarakan perasaan-perasaan lain. Beragamnya emosi dari masyarakat dalam merespon dinamika pemerintahan yang sedang berlangsung saat ini sudah berlangsung dari tahun ke tahun [4].

Emosi memiliki peran penting dalam proses komunikasi. Emosi didefinisikan sebagai keadaan mental yang disertai perubahan fisiologis pada individu [5]. Kondisi internal manusia dapat diungkapkan melalui berbagai macam bentuk komunikasi seperti ucapan/lisan, ekspresi wajah, dan tulisan [6]. Emosi yang disampaikan melalui teks sering kali mencerminkan perasaan terdalam seseorang. Selain itu, studi juga menunjukkan bahwa komunikasi yang mengandung unsur emosional memiliki pengaruh signifikan terhadap semua pikiran semua pihak yang terlibat [6]. Emosi manusia seperti marah, bahagia, sedih, dan takut yang tersirat dalam sebuah teks menjadi memperkuat pesan, mempengaruhi persepsi, dan mungkin mengubah sikap atau reaksi dari semua pihak yang terlibat atau berpartisipasi dalam interaksi tersebut. Akan tetapi, sering kali orang-orang masih menyalahartikan sebuah informasi karena mereka tidak memahami konteks emosional yang terkandung di dalamnya. Kurangnya pemahaman ini dapat menyebabkan kesalahpahaman dan konflik. Memahami dan menafsirkan emosi orang lain dapat menghasilkan interaksi sosial yang lebih baik, mengurangi konflik, dan meningkatkan kualitas hubungan sosial [7]. Terdapat banyak metode untuk dapat mengenali emosi dari sebuah teks, salah satunya adalah menggunakan bantuan machine learning.

Machine learning adalah salah satu cabang dari kecerdasan buatan yang berfokus pada pengembangan model yang memungkinkan komputer untuk membuat sebuah keputusan berdasarkan data yang diberikan [8]. *Supervised learning* merupakan salah satu pendekatan dalam machine learning yang sering digunakan untuk mengklasifikasikan sebuah data dari teks, salah satunya adalah emosi manusia. Dengan memanfaatkan *machine learning*, sebuah penelitian menunjukkan bahwa klasifikasi teks terbukti efektif karena mengurangi waktu dan tenaga yang diperlukan untuk pemrosesan secara manual dan memastikan keadilan dan objektivitas dalam penilaian [9]. Klasifikasi emosi dilakukan untuk memahami sentimen dan reaksi emosional masyarakat terhadap berbagai isu yang berkembang. Model algoritma *machine learning* seperti Support Vector Machine dan Random Forest sering digunakan untuk melakukan klasifikasi teks termasuk ekstraksi emosi.

Dalam penerapan *machine learning*, khususnya *Natural Language Processing*, memahami hubungan antarkata dianggap penting karena dapat meningkatkan performa model [10]. Banyak teknik yang dapat digunakan untuk memahami hubungan semantis antar kata dalam NLP, salah satu teknik yang cukup populer yaitu *word embedding*. *Word embedding* merupakan teknik yang digunakan untuk merepresentasikan kata dalam bentuk vektor agar dapat dipahami oleh algoritma *machine learning*. Penerapan *word embedding* dapat merepresentasikan mengenai hubungan antar kata dalam bentuk ruang vektor berkelanjutan [11].

Terdapat beberapa penelitian terdahulu yang mengangkat topik klasifikasi emosi dari teks dengan *word embedding*. Penelitian [12] yang berjudul “*Performance Comparison of TF-IDF And Word2Vec Models for Emotion Text Classification*” mendapatkan hasil skenario paling optimal untuk klasifikasi emosi menggunakan algoritma Support Vector Machine dengan pembobotan TF-IDF dengan nilai akurasi sebesar 93,45%, mengungguli algoritma Multinomial Naïve Bayes. Penelitian tersebut menggunakan dua tahap klasifikasi yaitu tahap pertama menentukan data mengandung emosi atau tidak, kemudian hasilnya digunakan untuk tahap selanjutnya mengategorikan emosi-emosi apa saja yang terkandung di dalamnya. Penelitian lainnya oleh [13] berjudul “*Performance Analysis of Different Word Embedding Models for Text Classification*” mendapatkan hasil berupa

algoritma Support Vector Machine dengan Word2Vec mengungguli skenario pemodelan lainnya. Word2Vec terbukti dapat menangkap hubungan semantis teks lebih baik daripada fitur ekstraksi lain seperti GloVe, Count Vectorizer, dan TF-IDF. Sementara itu, penelitian lain [14] berjudul “*Perbandingan Kinerja Word Embedding Word2Vec, Glove, Dan Fasttext Pada Klasifikasi Teks*” menunjukkan ketiga fitur *word embedding* yaitu Word2Vec, GloVe, dan FastText bekerja secara kompetitif. Namun, FastText menjadi fitur paling optimal dalam penelitian ini dengan nilai *F-measure* sebesar 0,979. Keunggulan FastText terletak pada kemampuannya dalam mengatasi masalah *out of vocabulary* atau vektor dari kata yang tidak ada pada korpus melalui representasi sub-kata.

Berdasarkan latar belakang yang telah diuraikan sebelumnya, penelitian ini akan mengaplikasikan algoritma *machine learning* Support Vector Machine dan Random Forest dengan tujuan untuk mengetahui perbandingan hasil performa kedua algoritma dengan teknik *word embedding* untuk mengklasifikasikan emosi pada komentar YouTube berbahasa indonesia mengenai topik kebijakan kenaikan Uang Kuliah Tunggal (UKT) dan kebijakan Tabungan Perumahan Rakyat. Penelitian ini akan menggunakan algoritma Support Vector Machine dan Random Forest dengan mengaplikasikan *word embedding*. SVM terbukti secara konsisten mendapatkan performa terbaik dalam mengklasifikasikan teks [15]. RF tangguh dalam mengatasi data dengan dimensi tinggi [16]. SVM dan RF menghasilkan performa yang cukup baik untuk melakukan tugas klasifikasi teks dengan data yang tidak seimbang [11]. Penelitian dimulai dengan pengumpulan data dari fenomena yang terjadi kemudian dilakukan proses pengolahan data, pelabelan data, klasifikasi, dan diakhiri dengan pengujian serta pengimplementasian model pada aplikasi yang akan dirancang. Proses klasifikasi akan membandingkan ekstraksi fitur *word embedding* dari dua algoritma *machine learning* yang digunakan untuk mendapatkan hubungan semantis teks yang terbaik. Hasil evaluasi model akan menggunakan beberapa metrik sebagai penilaian yaitu *classification report*, *confusion matrix*, *CPU usage*, dan *processing time*.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang dipaparkan sebelumnya maka dapat dirumuskan beberapa permasalahan seperti berikut:

1. Bagaimana hasil perbandingan performa model klasifikasi emosi dengan penggunaan teknik *word embedding* pada komentar pengguna YouTube terkait video kebijakan kenaikan Uang Kuliah Tunggal (UKT) dan kebijakan Tabungan Perumahan Rakyat.?
2. Bagaimana hasil visualisasi dari aplikasi untuk pengklasifikasian emosi terkait video kebijakan kenaikan Uang Kuliah Tunggal (UKT) dan kebijakan Tabungan Perumahan Rakyat menggunakan model yang paling optimal?

1.3 Batasan Masalah

Batasan masalah yang digunakan dalam skripsi ini adalah sebagai berikut:

1. Data yang digunakan dalam skripsi ini adalah komentar terkait kebijakan kenaikan Uang Kuliah Tunggal (UKT) dan kebijakan Tabungan Perumahan Rakyat.
2. Data yang digunakan dalam skripsi ini diambil dari YouTube menggunakan bantuan API YouTube. Data diambil dari kanal YouTube yang berkredibilitas seperti: METRO TV, Kompas.com, CNN Indonesia, CNBC Indonesia, dan Official iNews.
3. Rentang waktu pengambilan data yang digunakan dalam skripsi ini yaitu mulai dari tanggal 14 Februari 2024 hingga 14 Mei 2024.
4. Emosi diklasifikasikan menjadi enam class yaitu senang, sedih, marah, takut, jijik, dan kaget,

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan sebelumnya maka skripsi ini dibuat dengan tujuan sebagai berikut.

1. Mengetahui hasil komparasi performa model klasifikasi emosi dengan teknik *word embedding* pada komentar pengguna aplikasi YouTube pada konten video dengan topik kebijakan kenaikan Uang Kuliah Tunggal (UKT) dan kebijakan Tabungan Perumahan Rakyat.
2. Membuat sebuah sistem yang dapat mengklasifikasikan emosi dari sebuah teks.

1.5 Manfaat Penelitian

Skripsi ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Memberikan pemahaman yang lebih baik terhadap masyarakat bahwa machine learning dapat membantu mengenali konteks emosional yang terkandung dalam sebuah informasi berupa teks.
2. Dengan membandingkan algoritma, skripsi ini dapat mengidentifikasi kelebihan dan kekurangan setiap model dalam klasifikasi emosi berdasarkan fitur word embedding.
3. Menambah kontribusi terhadap ilmu pengetahuan khususnya terkait algoritma machine learning untuk klasifikasi emosi dan menjadi rujukan penelitian lanjutan atau pengembangan aplikasi serupa.