

1. Pendahuluan

1.1. Latar Belakang

Pemilihan presiden di Indonesia merupakan acara setiap lima tahun sekali yang dihadiri oleh seluruh rakyat Indonesia yang telah memiliki hak berdasarkan ketentuan peraturan perundang-undangan. Pemilihan umum telah diselenggarakan beberapa kali sepanjang sejarah Indonesia. Namun, pemilihan umum pertama yang diikuti langsung oleh rakyat Indonesia diselenggarakan pada era reformasi, tepatnya pada tahun 2004 [1]. Banyak isu yang mulai tersebar di media sosial hingga media berita baik *online* maupun cetak selama masa persiapan pemilihan presiden, yaitu kurang lebih 1 sampai 2 tahun menjelang pemilihan presiden. Berita dimuat di media berita dengan membuat judul yang menarik pembaca untuk membacanya. Namun, tidak ada filter untuk menentukan judul berita mana yang mengandung sentimen positif atau negatif [2]. Dengan begitu banyak berita, sulit untuk menentukan sentimen secara manual pada judul [3].

Media berita adalah media yang menyajikan informasi tentang peristiwa terkini, baik *online* maupun dalam bentuk cetak. Media berita *online* adalah informasi berita yang diakses secara *online* melalui *website* [3]. Media memiliki peran penting dalam sebuah ajang pemilu dengan menghadirkan berita-berita relevan yang dapat berdampak pada elektabilitas pasangan capres dan cawapres. Judul pada media berita *online* didesain menarik untuk memikat pembaca agar mengakses halaman berita dan membacanya.

Analisis sentimen adalah jenis teks yang melibatkan penggalian informasi dari sebuah teks untuk mengidentifikasi sentimen yang terkandung di dalamnya [4]. Analisis sentimen adalah cabang dari *Natural Language Processing* (NLP) yang menentukan apakah suatu teks mengandung sentimen positif atau negatif [5]. Data teks dapat diolah menggunakan analisis sentimen untuk mengetahui sentimen yang terkandung dalam sebuah judul berita. Karena banyaknya berita tentang pemilihan presiden 2024, maka penting untuk melakukan analisis sentimen terhadap judul berita.

1.2. Topik dan Batasannya

Untuk mengukur analisis sentimen dan membuat model klasifikasi, dataset yang digunakan adalah data teks judul berita yang diambil dari beberapa media berita *online* yang kemudian diberikan label 0 dan 1 yang merepresentasikan sentimen positif dan negatif dari judul berita tersebut. Penelitian ini dibangun dengan beberapa tahap preprocessing yakni *text cleaning*, *tokenization*, *stopword removal*, dan *stemming*. Tahap ekstraksi fitur dibangun dengan metode TF-IDF dan *FastText*. Tahap pembuatan model klasifikasi dibangun dengan hanya dibatasi pada algoritma *Support Vector Machine* (SVM).

1.3. Tujuan

Tujuan dari penelitian ini adalah membangun sistem analisis sentiment yang dapat mengklasifikasikan sentiment pada judul berita tentang Pemilihan Presiden Indonesia tahun 2024 yang dibangun dengan metode *Support Vector Machine* (SVM).