

1. Pendahuluan

Latar Belakang

Indonesia merupakan salah satu negara yang menganut sistem hukum demokrasi. Sistem demokrasi dibentuk agar masyarakat dapat berpartisipasi langsung untuk memilih kebijakan politik yakni menyuarakan pendapat yang nantinya dipakai untuk menentukan suatu keputusan. Pemilu adalah salah satu wujud dari sistem demokrasi ini. Sebelum teknologi berkembang dengan sangat pesat, masyarakat menyampaikan opini maupun kritik melalui media cetak sehingga para peneliti mengamati artikel-artikel yang dibuat oleh beberapa orang[1]. Hal ini berjalan dengan sangat baik karena pada masa tersebut media cetak memang merupakan salah satu media yang paling banyak digunakan oleh masyarakat untuk mendapatkan berbagai macam informasi secara detail. Seiring dengan perkembangan teknologi yang begitu pesat, media cetak sedikit demi sedikit mulai tersingkirkan. Berbagai macam media sosial bermunculan salah satunya yaitu twitter. Masyarakat tidak lagi membutuhkan media cetak untuk mengutarakan opini dan kritiknya secara panjang lebar. *Twitter* adalah sebuah media atau jaringan informasi yang diluncurkan pada tahun 2006. Masyarakat dapat mengungkapkan apa saja yang ingin disampaikan secara bebas dengan batasan 280 karakter pada setiap *tweet*. Karena kebebasan dalam menyampaikan pendapat terbuka luas, informasi informasi terbaru bisa dengan sangat cepat untuk didapatkan. Berdasarkan data dari statista[2], pada tahun 2018 pengguna *twitter* telah mencapai 350 juta orang dan terdapat 500 juta *tweet* setiap harinya. Maka dari itu, *twitter* merupakan wadah yang paling baik untuk pengambilan data dalam menganalisis sentimen masyarakat secara luas.

Menjelang pilpres tahun 2019, opini – opini atau *tweet* yang berkaitan dengan Presiden dan Calon Presiden adalah yang paling banyak diutarakan oleh pengguna *twitter* di Indonesia saat ini. Opini – opini masyarakat sangat penting untuk mengetahui keadaan keberpihakan masyarakat pada pemilihan presiden yang akan datang. Selain itu dengan melihat opini – opini yang terpampang secara bebas di *twitter*, kita dapat mengetahui secara umum keadaan beberapa aspek ekonomi, aspek kesehatan, dan aspek pembangunan dengan memanfaatkan opini – opini di *twitter* menggunakan metode *Naive Bayes Classifier* yang dikombinasikan dengan ekstraksi fitur *Lexicon SentiWordnet*.

Analisis sentimen masyarakat yang diklasifikasikan dalam beberapa aspek tersebut bermanfaat untuk mengetahui sejauh mana kinerja yang telah dilakukan oleh Presiden. Atas dasar ini, penulis bermaksud untuk melakukan analisis sentimen masyarakat dan mengklasifikasikannya kedalam aspek ekonomi, kesehatan, dan pembangunan. Sentimen masyarakat berkaitan dengan kinerja sangat penting untuk diketahui oleh seorang Presiden agar mengetahui tingkat kepuasan rakyatnya. Penelitian mengenai analisis sentimen telah banyak dilakukan sebelumnya dengan berbagai macam metode. Bahkan pada saat pemilihan Gubernur DKI Jakarta tahun 2017 juga telah dilakukan analisis sentimen oleh Buntoro[3]. Pada penelitian tersebut, mereka menggunakan kombinasi dari *Lexicon based* dengan *Naive Bayes* dan *Support Vector Machine* dan menyatakan bahwa *Lexicon Based* dapat mempengaruhi tingkat akurasi dari suatu program. Namun pada penelitian tersebut, Buntoro dan kawan - kawan hanya mengklasifikasikan dataset sebanyak 300 *tweet* terhadap dua kelas saja yaitu positif dan negatif dan mendapatkan akurasi sebesar 95% untuk metode *Naive Bayes*. Penelitian selanjutnya yang menggunakan *Lexicon based* yaitu Jea [4], Jea menggunakan *Lexicon Based* dalam tahapan ekstraksi fitur pada penelitiannya. *Dataset* yang digunakan yaitu *dataset* Bahasa Inggris dan metode klasifikasi yang digunakan yaitu *Support Vector Machine*.

Berdasarkan latar belakang yang sudah diuraikan, sistem ini dibuat dengan menerapkan metode *Naive Bayes* dalam mengklasifikasikan dataset *tweet* atau opini sebanyak 1508 data yang diambil dari *twitter* dan menggunakan *Lexicon SentiWordnet* untuk proses ekstraksi fitur. Penulis juga mencoba membandingkan hasil dari ekstraksi fitur *TF-IDF* dan menggabungkan *TF-IDF* dengan *Lexicon SentiWordnet*.

Tujuan dari penelitian ini adalah untuk mengetahui pengaruh penggunaan *Lexicon SentiWordnet* dalam tahap ekstraksi fitur yang kemudian akan di klasifikasi oleh metode *Naive Bayes* terhadap performansi sistem. Selanjutnya yaitu mengetahui pengaruh jika metode ekstraksi fitur dengan *TF-IDF* digabungkan dengan *Lexicon SentiWordnet*. Dan yang terakhir yaitu mengetahui tingkat kepuasan masyarakat terhadap kinerja Presiden dalam aspek ekonomi, kesehatan, dan pembangunan berdasarkan hasil klasifikasi oleh sistem.