

I. PENDAHULUAN

1.1. Latar Belakang

Analisis sentimen adalah metode yang digunakan untuk memahami maupun mengolah data untuk mendapatkan informasi berupa opini dalam bentuk teks dan diklasifikasikan sebagai opini positif atau negatif. Analisis sentimen bertujuan untuk mendalami pendapat ataupun opini pengguna berupa komentar di sosial media dengan tujuan untuk menunjukkan kesan maupun pengalaman pengguna terhadap suatu topik tertentu (Nasukawa & Yi, 2003). Analisis Sentimen adalah proses menganalisis, memahami pendapat, evaluasi, penilaian, sikap, dan emosi terhadap suatu entitas seperti produk, jasa, organisasi, individu, peristiwa, topik, secara otomatis untuk mendapatkan informasi (Liu, 2010).

Besarnya pengaruh dan manfaat dari Analisis Sentimen, menyebabkan penelitian ataupun aplikasi mengenai Analisis Sentimen berkembang pesat, bahkan di Amerika ada kurang lebih 20-30 perusahaan menggunakan Analisis Sentimen untuk mendapatkan informasi tentang sentimen masyarakat terhadap pelayanan perusahaan (SUMARTINI SARASWATI, 2011). Pada dasarnya Analisis Sentimen merupakan klasifikasi, tetapi kenyataannya tidak semudah proses klasifikasi biasa karena terkait penggunaan bahasa. Terdapat ambiguitas dalam penggunaan kata, tidak adanya intonasi dalam sebuah teks, dan perkembangan dari bahasa itu sendiri (Pang & Lee, 2008).

Untuk mengetahui metode terbaik yang dapat digunakan pada analisis sentimen maka dilakukan analisis perbandingan dari dua metode, ini dilakukan karena pada saat ini ada banyak penelitian mengenai analisis sentimen. Dalam melakukan analisis sentimen ada berbagai macam metode seperti *Naive Bayes*, *Maximum Entropy*, *Deep Learning LSTM*, *K-Nearest Neighbor (KNN)* dan *Support Vector Machine (SVM)*.

KNN merupakan salah satu dari metode algoritma *supervised* yang digunakan dalam mengklasifikasikan sebuah obyek dengan dengan membandingkan dengan mayoritas atribut dan training sample yang ada. Kelebihan KNN menggunakan prinsip yang sederhana, bekerja berdasarkan jarak terpendek dari sampel uji ke sampel latih dan tidak memperhitungkan kemungkinan distribusi dari masing-masing kelas. Sedangkan kekurangan dari metode KNN adalah *computational cost* dan sensitif terhadap derau dan input yang tidak sesuai karena KNN tidak memiliki weight atau bobot pada masing-masing atribut. Sedangkan SVM merupakan salah satu metode dalam *pattern recognition* yang bekerja atas prinsip menemukan *hyperplane* terbaik yang

memisahkan dua buah *class* pada input *space* dan mencari titik maksimalnya. SVM memiliki akurasi yang baik dalam pengenalan pola. Kelebihan lain dari metode SVM adalah pada proses *learning* yang cepat. Sedangkan kekurangan SVM sulit dipakai dalam problem jumlah sample besar. Perbandingan algoritma klasifikasi pernah dilakukan pada analisis sentimen terhadap perbankan, dengan K-NN mendapatkan 57% dan 49.62% di pengujian 100 dan 266 data, sedangkan SVM mendapatkan akurasi 63% dan 69.55% pada pengujian 100 dan 266 data (Probo et al., 2016). Pada tahun 2014 Bijalwan et al melakukan penelitian yang menghitung akurasi dari tiga metode klasifikasi yaitu KNN, *Term Graph* dan, *Naïve Bayes*. Penelitian tersebut menyatakan bahwa metode KNN memiliki akurasi tertinggi yaitu 98.95%. Berdasarkan hasil penelitian yang dilakukan peneliti lain mengenai komparasi dua metode salah satunya yang dilakukan oleh Hardian (2022) yang menggunakan dua metode yaitu *naive bayes* dan SVM mengenai Analisis Sentimen Pengguna *Twitter* Terhadap Kebijakan Kemdikbudristek Mengenai Kuota Internet Selama *Covid-19* membuktikan bahwa nilai akurasi SVM lebih tinggi dibandingkan *Naive Bayes*, dengan nilai akurasi SVM sebesar 97,% (Hardian, 2022). Penulis memilih menggunakan dua metode untuk dikomparasi, yaitu: *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM) dikarenakan dari beberapa data yang didapat dari penelitian sebelumnya menunjukkan bahwa metode KNN dan SVM merupakan metode yang memiliki tingkat akurasi yang tinggi.

Pada penelitian ini akan membandingkan metode *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM). *K-Nearest Neighbor* merupakan salah satu algoritma yang mudah diimplementasikan dengan tingkat keefektifan yang tinggi, serta cocok untuk berbagai masalah yang berhubungan dengan klasifikasi. Sedangkan *Support Vector Machine* merupakan salah satu algoritma yang tepat dipakai untuk klasifikasi *text classification*. Kemampuan *Support Vector Machine* menemukan hyperplane terbaik menjadikan algoritma ini memiliki tingkat generalitas yang tinggi serta menjadikannya algoritma dengan akurasi terbaik dibandingkan algoritma lainnya (Marimuthu et al., 2019). Perbandingan ini bertujuan untuk mengetahui algoritma mana yang memiliki akurasi terbaik dan waktu proses tercepat analisis perbandingan dengan metode *K-Nearest Neighbor* dan *Support Vector Machine*. Sehingga nantinya dapat dihasilkan suatu kesimpulan mengenai metode yang memiliki tingkat efektifitas dan efisiensi terbaik yang nantinya juga dapat menjadi rekomendasi bagi peneliti lainnya.

Dalam penelitian ini, data sampel yang digunakan adalah data komentar masyarakat terhadap kuliah luring pasca kebijakan *new normal*. *New Normal* merupakan kehidupan baru di mana masyarakat tetap melakukan berbagai

aktivitas seperti biasa namun tetap menerapkan protokol kesehatan yang telah ditetapkan pemerintah agar penyebaran Virus *Covid-19* dapat teratasi. Seiring dengan berjalannya kebijakan *New Normal* masyarakat (pelajar) perlahan mulai diarahkan untuk mengikuti pembelajaran atau kuliah secara luring kembali. Pembelajaran luring dapat diartikan sebagai pembelajaran yang dilakukan di luar tatap muka oleh guru dan peserta didik, namun dilakukan secara offline yang berarti guru memberikan materi berupa tugas hardcopy kepada peserta didik kemudian dilaksanakan di luar sekolah maupun perkuliahan (Kemendikbud, 2020).

Kebijakan *New Normal* ini menimbulkan beberapa pendapat di masyarakat, ada yang setuju dengan diberlakukannya pembelajaran luring dan ada yang tidak setuju. Persepsi siswa terhadap pembelajaran kelas secara luring dengan adaptasi kebiasaan baru adalah positif atau dapat dikatakan sangat setuju. karena lebih mudah memahami materi di kelas, dapat melakukan tanya jawab materi pelajaran secara langsung, selain itu siswa juga menikmati proses pengerjaan tugas-tugas yang diberikan di kelas secara langsung bersama guru. (Sartinem, 2021). Dalam pembelajaran luring pada masa pandemi *Covid-19* terdapat beberapa kendala seperti terbatasnya waktu pembelajaran karena waktu pembelajaran di kurangi sehingga materi yang di sampaikan tidak tuntas (Risdianto et al., 2020).

Untuk mengetahui tanggapan masyarakat dalam pelaksanaan pembelajaran secara luring ataupun opini masyarakat tentang pembelajaran secara luring, terdapat media sosial yang dapat dipergunakan. Salah satunya yang populer pada saat ini yaitu *Twitter*. *Twitter* adalah situs mikroblogging yang sebenarnya dikembangkan untuk telepon seluler, dirancang agar pengguna memposting sesuatu yang singkat, teks sebanyak 140 karakter. *Twitter* memiliki model pertemanan langsung yang mana pengguna bebas memilih akun *Twitter* siapa yang mereka ikuti (follow) dan mereka juga bisa menentukan kelompok pengikut (followers) (Marwick & Boyd, 2011).

Berdasarkan latar belakang diatas, pada penelitian ini dilakukan analisis sentimen untuk mengetahui perbandingan antara metode *K-Nearest Neighbor* dan *Support Vector Machine*. penelliti akan fokus membahas mengenai metode mana yang paling baik digunakan dalam pengimplementasian analisis sentimen di masyarakat atau pengguna *Twitter* mengenai kebijakan *new normal* terhadap kuliah luring. Maka, dari penjelasan tersebut peneliti tertarik untuk membahas terkait “PERBANDINGAN METODE *K-NEAREST NEIGHBOR* DAN *SUPPORT VECTOR MACHINE* DALAM ANALISIS SENTIMEN MASYARAKAT TERHADAP KULIAH LURING PASCA KEBIJAKAN *NEW NORMAL*)”.

1.2. Rumusan Masalah Penelitian

Berdasarkan latar belakang diatas, maka dapat dirumuskan beberapa permasalahan pada penelitian kali ini yaitu sebagai berikut :

1. Bagaimana perbandingan tingkat akurasi yang didapatkan dari hasil klasifikasi sentimen masyarakat terhadap kuliah luring pasca kebijakan *new normal* menggunakan metode *K-Nearest Neighbor* dan *Support Vector Machine*?
2. Bagaimana arah sentimen masyarakat terhadap kuliah luring pasca kebijakan *new normal* menggunakan metode *K-Nearest Neighbor* dan *Support Vector Machine* ?

1.3. Tujuan Penelitian

Adapun tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut :

1. Mengetahui perbandingan tingkat akurasi yang didapatkan dari analisis sentimen dengan metode *K-Nearest Neighbor* dan *Support Vector Machine*.
2. Mengetahui hasil dari arah sentimen masyarakat terhadap kuliah luring pasca kebijakan *new normal* menggunakan metode *K-Nearest Neighbor* dan *Support Vector Machine*.

1.4. Manfaat Penelitian

Hasil penelitian ini diharapkan bermanfaat untuk:

1. Hasil komparasi metode analisis sentimen dapat dijadikan sebagai acuan bahan referensi bagi peneliti serupa dan bahan evaluasi terhadap pelaksanaan analisis sentimen yang akan datang.
2. Memberikan informasi arah sentimen masyarakat terhadap kuliah luring pasca kebijakan *new normal* berupa sentimen positif, negatif dan netral.

1.5. Batasan Penelitian

Batasan masalah dalam penelitian ini adalah :

1. Dalam penelitian ini, sampel yang digunakan untuk melakukan uji klasifikasi sentimen digunakan data: tanggapan masyarakat terhadap kebijakan *New Normal* yang diperoleh melalui *Twitter*
2. Penelitian ini tidak terfokus pada alasan yang mempengaruhi pendapat masyarakat terhadap kebijakan pasca *new normal*, melainkan terfokus pada keakurasian dalam klasifikasi yang dilakukan oleh kedua metode/algoritma
3. Hasil akhir dalam penelitian ini adalah informasi mengenai metode analisis sentimen terbaik yang dapat digunakan untuk kondisi tertentu (pada kasus ini, yaitu: sentimen masyarakat terhadap kuliah luring pasca *new normal*).