

1. Pendahuluan

Latar Belakang

Media sosial merupakan sarana untuk melakukan aktifitas secara online seperti komunikasi, promosi, atau membuat konten. Twitter merupakan media sosial yang sudah dipakai sejak tahun 2006 dan memiliki cukup banyak pengguna yaitu sekitar ± 18 juta di Indonesia. Twitter merupakan media sosial yang populer di Indonesia, dengan kepopuleran tersebut media ini digunakan untuk sarana promosi, kampanye, dan berbagi informasi. Banyaknya pengguna pada twitter inilah yang mendorong penelitian ini menggunakan twitter sebagai sumber penelitian. Oleh karena itu data yang digunakan untuk penelitian akan diambil dari twitter API menggunakan `snsrape python` dengan batasan waktu Juni 2020 sampai November 2022 dan topik yang diambil hanya mengenai topik bank BCA.

Pada penelitian yang dilakukan oleh Indriyana[1] pada tahun 2019 menunjukkan bahwa penggunaan metode BM25 dan pengklasifikasian menggunakan `imporved K-NN` yang digunakan untuk analisis sentimen pada ulasan Mobile BCA mendapatkan hasil akurasi yang fluktuatif, karena dipengaruhi besarnya `k-values`. Pada penelitian yang dilakukan oleh Ferly[2] pada tahun 2017 menunjukkan bahwa analisis ulasan mengenai aplikasi Mobile BCA diperoleh bahwa `Naïve Bayes Levenshtein Distance` memiliki akurasi yang lebih besar dibandingkan `Naïve Bayes tanpa Levenshtein Distance`. Pada penelitian yang dilakukan oleh Muhammad asjad[3] pada tahun 2021 menunjukkan bahwa analisis ulasan pada sebuah film pada metode `Random Forest` yang menggunakan `skipgram Word2Vec 300` dimensi yang kemudian diterapkan `Adaptive Boosting` mendapat akurasi sebesar 75,76%.

Pada penelitian yang dilakukan Lusiana[4] pada tahun 2018 mengenai analisis sentimen dengan menggunakan metode SVM menunjukkan bahwa dari data yang diklasifikasi dibagi menjadi dua bagian yaitu secara manual dan otomatis menggunakan model classifier. Data yang diklasifikasi secara manual dibagi menjadi data *testing* dan data *training* dengan rasio 2:8. Kernel yang digunakan pada penelitian yaitu RBF untuk mengetahui tingkat akurasi. Pada penelitian yang dilakukan oleh Anita[5] pada tahun 2015, penggunaan metode SVM pada analisis opini masyarakat di tweet mengenai kemacetan di kota besar di Indonesia menunjukkan variabel yang berpengaruh yaitu jumlah data, perbandingan data latih dan uji, dan perbandingan jumlah data positif dan negatif. Pada penelitian yang dilakukan oleh Herlina[6] pada tahun 2018 pada analisis pada pemilihan studi lanjut dengan membandingkan kernel SVM, diperoleh akurasi tertinggi pada kernel *Polynomial* dibandingkan dengan kernel *Linear* dan RBF.

Setiap pengguna media sosial menggunakan media tersebut untuk mengutarakan setiap pendapat, pemikiran, membuat konten maupun sebagai sarana iklan. Dan ketika mereka membuat komentar yaitu berupa tweet, maka setiap kata yang mereka tulis mengandung nilai sentimen yang bisa mempengaruhi nama baik suatu kasus, brand, maupun perusahaan. Dalam kasus ini penulis memilih untuk melakukan analisis sentimen sebagai upaya untuk mengetahui nilai sentimen dari opini yang diberikan oleh masyarakat di Twitter. Pada penelitian ini penulis akan menggunakan opini pengguna Bank BCA karena pada periode Juni 2020 sampai November 2022 ramai dibicarakan di twitter mengenai gangguan pada mobile banking BCA, selain itu dengan dilakukannya penelitian ini masyarakat bisa mengetahui nilai terhadap Bank BCA dan juga bisa membantu masyarakat untuk menilai dan memilih perbankan yang akan mereka gunakan.

Analisis sentimen merupakan cara yang dapat digunakan untuk mengetahui kualitas suatu topik di masyarakat. Terdapat tiga nilai sentimen yaitu nilai positif, negatif, dan netral. Dan setiap perkataan mengandung nilai sentimen. Penelitian ini akan menggunakan data yang diambil dari twitter yang mengutarakan pendapatnya tentang Bank BCA. Penelitian ini menggunakan metode *Support Vector Machine (SVM)*, SVM merupakan algoritma Machine Learning yang dapat digunakan untuk klasifikasi text. Penelitian ini akan dengan membandingkan 3 kernel yaitu *Radial Basis Function*, *Linear*, dan *Polynomial*. Penelitian ini dilakukan agar bisa mengetahui tingkat akurasi kernel SVM bila digunakan pada klasifikasi sentimen. Jika akurasi yang didapat tinggi pada kernel SVM yang digunakan untuk analisis sentimen maka hasil yang didapat juga akurat.

Tujuan analisis ini adalah melakukan perbandingan menggunakan metode kernel *Support Vector Machine (SVM)* untuk mengetahui tingkat akurasi yang paling tinggi dari model kernel yang dibandingkan dan mengetahui seberapa baik kinerja metode SVM pada kasus analisis sentimen. Peneliti memilih membandingkan kernel dikarenakan dengan membandingkan kernel sehingga dapat mengetahui tingkat akurasi kernel terbaik dalam kasus klasifikasi sentimen, selain itu perbandingan ini juga dilakukan karena dari studi literatur yang telah dilakukan belum ada perbandingan penggunaan kernel dalam studi kasus analisis sentimen, sehingga peneliti yang menggunakan metode SVM untuk analisis sentimen tidak mendapat hasil akurasi yang optimal karena penggunaan kernel yang tidak tepat. Terdapat paper yang melakukan perbandingan kernel namun bukan pada studi kasus analisis sentimen. Penelitian ini sangat penting untuk dilakukan karena untuk mengetahui tingkat akurasi kernel SVM terbaik dalam klasifikasi sentimen, sehingga apabila ingin melakukan analisis sentimen

menggunakan kernel SVM sudah diketahui kernel terbaik yang dapat digunakan. Kalimat sentimen yang ada akan dibagikan menjadi tiga kelompok yaitu positif, negatif, dan netral lalu akan diklasifikasi menggunakan metode *Support Vector Machine (SVM)*.

Topik dan Batasannya

Penelitian ini adalah dengan membangun sistem analisis sentimen menggunakan metode kernel SVM. Sistem ini direkomendasikan bagi pengguna laptop atau PC, karena sistem ini dibangun dengan memanfaatkan google colab. Penggunaan google colab bertujuan untuk mempermudah seseorang untuk menggunakan sistem ini. Sistem akan meminta input berupa data csv sebagai dataset. Sistem ini memiliki batasan yaitu dataset yang diinput juga memiliki batasan waktu dan diperuntukkan untuk data berupa teks dan memiliki label. Apabila tidak memiliki label bisa melakukan pelabelan manual maupun menggunakan analisis pelabelan. Selanjutnya sistem akan mengolah data dengan pemodelan yang telah dibangun. Output yang muncul akan berupa hasil akurasi, *classification report*, dan diagram bar.

Penulis melakukan perbandingan pada kernel SVM yaitu *Radial Basis Function*, *Linear*, dan *Polynomial*. Hal ini dilakukan untuk mengetahui klasifikasi kernel SVM yang paling baik.

Tujuan

Tujuan dari penelitian ini adalah melakukan perbandingan menggunakan metode kernel *Support Vector Machine (SVM)* untuk mengetahui tingkat akurasi yang paling tinggi dari model kernel yang dibandingkan dan mengetahui kinerja metode SVM pada kasus analisis sentimen. Kalimat sentimen yang ada akan dibagikan menjadi tiga kelompok yaitu positif, negatif, dan netral menggunakan metode *Support Vector Machine (SVM)*. Sistem ini dibangun dengan maksud untuk mempermudah suatu lembaga/perusahaan dalam mengetahui nama baik mereka dikalangan masyarakat.

Tabel 1 Tujuan, Pengujian dan Kesimpulan

No	Tujuan	Pengujian	Kesimpulan
1	Mengetahui kinerja metode SVM pada kasus analisis sentimen	Dilakukan klasifikasi dengan model SVM sehingga mengetahui tingkat akurasi model	Mengetahui tingkat akurasi kernel SVM pada kasus analisis sentiment dengan baik
2	Melakukan perbandingan menggunakan metode kernel <i>Support Vector Machine (SVM)</i>	Dilakukan pengujian 3 kernel dengan 3 kali percobaan dengan menggunakan parameter yang berbeda disetiap percobaan nya	Percobaan ini dilakukan untuk mengetahui kernel terbaik dengan parameter terbaik dalam kasus klasifikasi nilai .sentimen

Organisasi Tulisan

Rencana kegiatan yang akan dilakukan dalam analisis ini, yaitu :

1. Analisa masalah dan solusi
 Pada tahap ini dilakukan analisis pada opini dan komentar di twitter, lalu melakukan studi literatur terkait metode, algoritma dan cara penerapannya. Selanjutnya menentukan solusi dan Batasan masalah yang akan diteliti.
2. Pengumpulan data
 Pada tahap ini akan dilakukan pengumpulan data yang akan digunakan untuk analisis.
3. Analisa data
 Pada tahap ini dilakukan analisis pada data yang telah dikumpulkan agar bisa mengetahui hipotesis sementara dari data yang dikumpulkan.
4. Perancangan sistem

Pada tahap ini dilakukan perancangan sistem analisis sentimen menggunakan metode kernel SVM. Dalam tahap ini juga dilakukan pembersihan data agar seluruh data menjadi data yang relevan, *case folding*, penghapusan stopwords, dan *stemming*.

5. *Testing*

Pada tahap ini dilakukan *testing* pada data yang telah disiapkan dan dianalisis menggunakan metode kernel SVM.

6. Evaluasi

Pada tahap ini dilakukan evaluasi atau kesimpulan pada hasil data yang sudah diolah menggunakan metode kernel SVM.