

1. Pendahuluan

Latar Belakang

Adanya konflik menyebabkan harga minyak dunia tidak stabil karena Rusia adalah salah satu produsen minyak terkemuka di dunia. Pemulihan ekonomi global juga mungkin lebih lemah dari perkiraan sebelumnya. Pemulihan ekonomi global pasca-COVID-19 dengan ancaman inflasi sudah terlihat di beberapa negara maju, mulai dari Amerika Serikat hingga Indonesia. Kenaikan harga bahan bakar minyak (BBM) berdampak pada tambahan inflasi sebesar 1,9%. Dengan kombinasi tersebut, bahan pangan terjangkau terjaga sehingga inflasi dapat dijaga di bawah 7% hingga akhir tahun, sedangkan di negara maju, kenaikan harga energi diperkirakan akan menurunkan pertumbuhan ekonomi sebesar 0,69 poin proporsi di tahun 2022 dan 0,95 poin proporsi di tahun 2023[1]. Untuk mengetahui respon masyarakat terhadap kenaikan harga BBM, khususnya pada aspek ekonomi, metode yang dapat digunakan untuk mengetahui respon masyarakat tersebut adalah analisis sentimen.

Analisis sentimen adalah cara untuk menilai opini tertulis atau lisan untuk menentukan apakah sebuah opini bersifat positif, negatif, atau netral[2]. Salah satu metode atau algoritma yang populer untuk mengklasifikasikan analisis sentimen adalah Algoritma Naïve Bayes[3], seperti yang dilakukan oleh Saputra dalam penelitiannya mengenai topik kenaikan harga BBM[4]. Namun Algoritma SVM juga dapat digunakan untuk analisis sentimen[5].

Beberapa ekstraksi fitur sudah sering ditambahkan dalam analisis sentimen[6] dengan menganalisis sentimen mengenai calon Gubernur DKI Jakarta 2017 menggunakan Lexicon dengan Naive Bayes dan SVM. Salah satu peneliti yang menggunakan fitur ekstraksi leksikon dalam penelitiannya[7] kemudian ada peneliti yang menggunakan Lexicon Inset dalam penelitiannya[8]. Tanda baca dan huruf kapital pada kalimat perlu dihilangkan dengan menggunakan Cleansing atau Case Folding dan merubah setiap kata pada kalimat menjadi kata dasar dengan menggunakan metode Stemming dan mengetahui tingkat akurasi yang didapat dari hasil klasifikasi menggunakan algoritma SVM, preprocessing data merupakan hal yang sangat penting dalam analisis sentimen karena dapat meningkatkan akurasi[9].

Data yang digunakan dalam penelitian ini diperoleh dari Twitter karena Twitter merupakan salah satu media sosial terbesar di dunia yang menyaingi Facebook[10], yang kini telah berganti nama menjadi X untuk mendukung bisnis yang akan dikembangkan oleh perusahaan milik Elon Musk tersebut[11]. Penelitian ini berfokus pada perbandingan performa klasifikasi SVM dengan ekstraksi fitur TF-IDF dan performa klasifikasi Lexicon Inset yang telah dibuatkan kamus untuk setiap kata[12], dari perbandingan tersebut akan terlihat akurasi dari klasifikasi yang dilakukan.

Topik dan Batasannya

Topik pada penelitian ini membahas mengenai klasifikasi sentimen menggunakan Lexicon Inset dan metode Support Vector Machine(SVM) dan menerapkan K-Fold Cross Validation untuk mengukur kinerja suatu model atau algoritma machine learning dengan lebih akurat. Penelitian ini akan dilakukan analisis berdasarkan nilai akurasi dan F1-Score.

Batasan masalah pada penelitian ini adalah menggunakan dataset berupa *tweet* dari aplikasi Twitter dengan menggunakan *keyword* “BBM naik” dan hanya menggunakan label positif dan negatif, tidak menggunakan label netral dikarenakan jika menggunakan label netral terjadi imbalance data yang sangat ekstrim dan bisa mempengaruhi performa sistem.

Tujuan

Tujuan pada penelitian ini adalah melakukan perbandingan klasifikasi sentimen berdasarkan tweet di Twitter mengenai kenaikan harga BBM menggunakan metode SVM dan Lexicon. Hasil penelitian ini akan menunjukkan nilai *accuracy* dan *F1-Score*.