Kemunculan berbagai jenis aplikasi media sosial baru merupakan hasil dari perkembangan teknologi informasi dan komunikasi. Media sosial baru mendorong kemampuan individu manusia dalam menerima, memproses, dan menyebarkan informasi yang mereka dapatkan. Media sosial baru dapat diartikan sebagai media yang terbentuk dari interaksi antara manusia dan komputer yang terhubung melalui internet (Fardiah, Darmawan, dan Chatamallah, 2014). Kemunculan media sosial baru ini membawa perubahan signifikan dalam kehidupan manusia, memengaruhi pola kehidupan masyarakat, budaya, cara berpikir, dan hampir semua aspek kehidupan manusia. Salah satu media sosial yang sangat populer saat ini, baik di seluruh dunia maupun di Indonesia, terutama di kalangan mahasiswa, adalah Twitter.Twitter adalah media sosial gratis dan terpopuler serta menyediakan layanan jaringan yang memungkinkan pengguna untuk berbagi pendapat melalui pesan singkat atau sering dikenal dengan tweet [1].

Analisis sentimen merupakan cara mengumpulkan pendapat khalayak umum menggunakan jejaring sosial yang didalamnya terdapat mengandung pelayanan umum serta isu terkini [2]. Analisis Sentimen adalah mengekstrak opini public tentang topik tertentu, produk atau jasa yang di dalam nya terkandung teks-teks yang tidak terstruktur [3]. Tujuan dari analisis sentimen merupakan untuk mengetahui opini seseorang ke arah positif atau negatif yang nantinya dapat digunakan untuk pengambilan keputusan.

D. K. Zala, “*A Review on Basic Methodology of Twitter Base Prediction System*,” pp. 447–451, 2018

M. Syarifuddin, “*Analisis Sentimen Opini Publik Mengenai Covid-19 pada Twitter Menggunakan Metode Naïve Bayes dan Knn*,” *Inti Nusa Mandiri*, Vol. 15, No. 1, pp. 23– 28, 2020.

S. Saidah and J. Mayary, “*Analisis Sentimen Pengguna Twitter Terhadap Dompet Elektronik Dengan Metode Lexicon Based dan K – Nearest Neighbor,” J. Ilm. Inform. Komput*., Vol. 25, No. 1, pp. 1–17, 2020, doi: 10.35760/ik.2020.v25i1.2411.

Generasi Z merupakan generasi yang lahir antara tahun 1995 hingga 2010. Sebagian di antara mereka memasuki fase perkembangan remaja dan dewasa awal (Santrock, 2018). Generasi Z tumbuh setelah menjamahnya internet (Pichler, 2021), di mana dunia digital yang maju begitu pesatnya (Gentina, 2020). Menurut Roberts dkk, (2014), remaja dari generasi Z menjadikan smartphone sebagai bagian dari kehidupan sehari-hari. Efeknya, remaja menginternalisasi nilai dalam hidup dari berbagai agen sosialisasi(Berns, 2013). Generasi merupakan kelompok individu yang mengalami konteks sejarah dan sosiokultural yang sama pada tahap perkembangan tertentu (Lyons & Kuron, 2013). Genarasi Z ialah kategori kelompok dengan jumlah paling dominan dari berbagai kategorisasi generasi baik di Indonesia dengan 27,94% populasi (BPS, 2020), maupun di berbagai belahan dunia (Spitznagel, 2020). Generasi Z dikenal sebagai generasi digital native yang telah mengenal teknologi seperti gawai sejak kecil (Roberts dkk, 2014). Generasi ini juga mampu mengekspresikan identitas mereka baik di lingkungan nyata maupun maya. Remaja generasi Z menerima sosialisasi nilai dari berbagai pihak lantaran kecanggihan dunia digital. Secara singkat, sosialisasi dapat dimaknai sebagai proses di mana individu menyesuaikan diri dan menginternalisasi nilai, norma, kebiasaan dan perilaku dari kelompok sosial (Perez-Falkner, 2013).Penelitan Broadbent dkk (2017) menunjukkan bahwa nilai-nilai hidup generasi Z disosialisasikan dan dipengaruhi oleh orangtua (89%), teman (78%), guru (70%), selebriti (30%), dan politisi (17%). Penelitian tersebut menegaskan urgensi keluarga, khususnya orangtua, dalam fungsi sosialisasi, di samping keempat fungsi yang lain lain yakni reproduksi, dukungan ekonomi, dukungan sosial, dan peran sosial (Lestari, 2016).

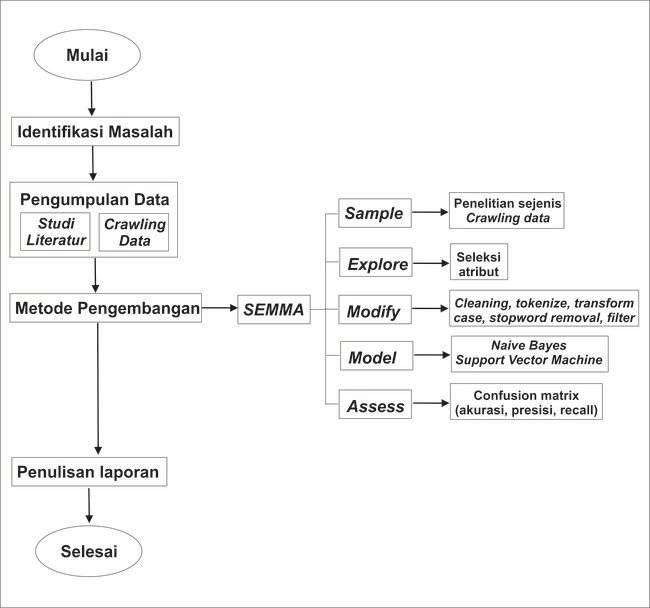
Metode dalam data mining adalah metode Knowledge Discovery in Database (KDD), metode Cross Industry Standard Process (CRISP-DM) dan metode SEMMA. Setiap metode mempunyai kelebihannya masing-masing yang dapat diterapkan sesuai dengan analisis masalah dan hasil yang diharapkan (Saputra, 2022).

Data set harus melalui tahap text preprocessing terlebih dahulu karena data set tidak bisa digunakan tanpa melalui tahap pengelolaan data. Text preprocessing adalah suatu proses pengelolaan data set sebelum data tersebut diproses. Pada kenyataannya, masih banyak data set yang tidak bersih seperti kesalahan sistem saat pencatatatan sehingga terjadinya data duplikat. Data yang belum diolah atau data tidak bersih kategorinya seperti format data yang tidak beraturan, adanya data kosong, tipe data yang berbeda-beda, adanya atribut yang tidak penting, dan lain sebagainya. Semakin bersih pra proses yang dilakukan, maka kemungkinan besar hasil data tersebut semakin akurat (Irwansyah Saputra, 2022).

Metode Support Vector Machine (SVM) termasuk jenis machine learning supervised learning yang membutuhkan sample data. Support Vector Machine (SVM) merupakan algoritma yang dikembangkan oleh Boser, Guyon, dan Vapnik pada tahun 1992. Support Vector Machine (SVM) memiliki konsep yang dikombinasi dari teori komputasi sebelumnya. Metode ini dapat mengubah data latih ke dimensi yang lebih tinggi menggunakan pola non linear (Hilda Kusumahadi et al., 2019). Metode ini juga sering digunakan dalam klasifikasi yang termasuk ke dalam supervised learning dan mempunyai konsep matematis yang jelas. Jika dibandingkan dengan metode lain, metode ini mempunyai tingkat akurasi yang lebih tinggi (Pratiwi et al., 2021). Support Vector Machine (SVM) dapat bekerja pada data set yang mempunyai dimensi tinggi menggunakan kernek trik (Pane et al., 2021)

Confusion matrix adalah matriks yang digunakan untuk melakukan evaluasi proses model klasifikasi berupa jumlah data uji yang benar dan salah. Dengan adanya matriks ini dapat mengetahui kualitas kinerja model klasifikasi (Normawati & Prayogi, 2021). Matriks ini berisi data target prediksi yang 21 dibandingkan dengan data target aktual. Data prediksi merupakan nilai yang didapatkan dari hasil pemodelan machine learning, sedangkan data aktual adalah nilai sebenarnya yang dimiliki.

Mengetahui analisis sentimen pada media sosal twitter terhadap topik tentang Generasi Z menggunakan metode Support Vector Machine (SVM). Dan Mengetahui sentimen positif atau sentimen negatif pada data set twitter mengenai Generasi Z.



Studi Literatur

Crawling Data

Pengumpulan Data

Identifikasi Masalah

Metode Pengembangan

SEMMA

Modify

Explore

Sample

Assess

Model