

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Penelitian sentimen analisis menggunakan *Support Vector Machine* sudah ada di beberapa jurnal/skripsi, sudah ada penelitian tentang penggunaan *Support Vector Machine* sebagai metode dalam melakukan analisis sentimen. Beberapa diantara penelitian/pustaka yang relevan dengan topik yang dibawakan sebagai berikut:

Table 2.1 Tinjauan Pustaka

No.	Nama Penulis	Objek	Teknologi	Hasil
1.	Rian Tineges, Agung Triayudi, Ira Diana Sholihati	Opini masyarakat terhadap layanan IndiHome berdasarkan <i>Twitter</i>	Klasifikasi <i>Support Vector Machine (SVM)</i>	Hasil pengujian didapat akurasi sebesar 87% dengan <i>precision</i> sebesar 86%, <i>recall</i> sebesar 95%, <i>error rate</i> sebesar 13%, dan untuk nilai perbandingan rata-rata <i>precision</i> dan <i>recall (f1-score)</i> adalah sebesar 90%.
2.	Wanda Athira Luqyana	<i>Cyberbullying</i> pada komentar <i>Instagram</i>	<i>Term Frequency-Inverse Document Frequency (TF-IDF)</i> dan metode klasifikasi <i>Support Vector Machine (SVM)</i>	Hasil pengujian didapatkan hasil akurasi tertinggi sebesar 90% pada komposisi data latih 50% dan komposisi data uji 50%.

No.	Nama Penulis	Objek	Teknologi	Hasil
3.	Ferdi Alvianda	Konten radikal di media sosial <i>Twitter</i>	<i>Support Vector Machine (SVM)</i>	Tingkat akurasi tertinggi yang dihasilkan dari penelitian ini adalah sebesar 70%.
4.	Fadholi Fat Haranto, Bety Wulan Sari	Opini pengguna <i>Twitter</i> terhadap pelayanan Telkom dan Biznet	<i>Support Vector Machine (SVM)</i>	Hasil analisis menghasilkan <i>accuracy</i> 79,6%, <i>precision</i> 76,5%, <i>recall</i> 72,8%, dan <i>f1-score</i> 74,6% hasil dari Telkom, dan nilai <i>accuracy</i> 83,2%, <i>precision</i> 78,8%, <i>recall</i> 71,6%, dan <i>f1-score</i> 75% hasil dari Biznet.
5.	Fadhilah Dwi Ananda, Yoga Pristyanto	Komentar pengguna <i>Twitter</i> terhadap layanan <i>Internet Service Provider</i>	<i>Support Vector Machine (SVM)</i>	Hasil pengujian didapatkan akurasi sebesar 90%, <i>precision</i> 88,63%, <i>recall</i> 88,63%
6.	Tediyan Rahmat Wibowo	Sentimen Analisis Keluhan Pelanggan Menggunakan Metode <i>Support Vector Machine (SVM)</i>	<i>Support Vector Machine (SVM)</i>	Hasil akurasi pengujian yang dihasilkan. Sentimen yang dihasilkan dari keluhan pelanggan menggunakan model yang sudah dibuat.

Tabel di atas menyajikan penelitian yang sudah dilakukan terlebih dahulu mengenai topik tentang analisis sentimen menggunakan metode *Support Vector Machine* dengan output tingkat akurasi hasil pemodelan. Informasi yang dipaparkan menjadi referensi bagi penulis dalam melakukan penelitian ini.

Dalam penelitian ini output yang dihasilkan yaitu tingkat akurasi dari hasil pemodelan dengan penambahan implementasi hasil model kedalam *API* sehingga bisa digunakan pada aplikasi keluhan pelanggan.

2.2 Dasar Teori

Dasar teori yang digunakan dalam penelitian meliputi analisis sentiment, *Support Vector Machine (SVM)*, *text preprocessing*, evaluasi, *Word Cloud*.

2.2.1 Analisis Sentimen

Analisis sentimen atau *opinion mining* merupakan bidang studi yang menganalisis opini, sentimen, evaluasi, penilaian, sikap dan emosi seseorang terhadap sebuah produk, layanan, organisasi, individu, masalah, peristiwa atau topik. Analisis sentimen atau *opinion mining* berfokus untuk melihat pendapat atau kecenderungan opini terhadap suatu masalah atau objek oleh seseorang, apakah beropini positif atau negatif. Salah satu dari kegunaannya adalah untuk membantu mengetahui pendapat dari masyarakat atau seseorang terhadap suatu produk atau jasa dengan menggunakannya sebagai alat bantu untuk melihat pendapat konsumen terhadap produk atau jasa tersebut. Analisis sentimen sangat bermanfaat dalam dunia usaha seperti melakukan analisis terhadap suatu produk atau jasa untuk melihat respon konsumen terhadap produk atau jasa tersebut, sehingga dapat membuat strategi-strategi baru pada tahapan-tahapan berikutnya. (Fadhilah Dwi Ananda, Yoga Pristyanto, 2021)

2.2.2 *Support Vector Machine (SVM)*

Support Vector Machine (SVM) merupakan salah satu metode dalam supervised learning yang biasanya digunakan untuk klasifikasi seperti (*Support*

Vector Classification) dan regresi (*Support Vector Regression*). Dalam pemodelan klasifikasi, *SVM* memiliki konsep yang lebih matang dan lebih jelas secara matematis dibandingkan dengan teknik-teknik klasifikasi lainnya. *SVM* juga dapat mengatasi masalah klasifikasi dan regresi dengan linear maupun non linear. (Samsudiney, 2019)

Table 2.2 Algoritma SVM
Algoritma Support Vector Machine

<p>1. Kelas data: $y_i \in \{-1, +1\}$, untuk $i = 1, 2, \dots, i$.</p> <p>2. Diperoleh decision function berikut :</p> $f(x) = \text{sign}(w \cdot x + b)$ <p>3. Berdasarkan pada <i>decision function</i>, diasumsikan kedua kelas dapat terpisah oleh <i>hyperplane</i> berdimensi d:</p> <p>a. $w^{\rightarrow} \cdot x + b = 0$</p> <p>b. $w^{\rightarrow} \cdot x + b \leq -1$</p> <p>c. $w^{\rightarrow} \cdot x + b \geq +1$</p> <p>4. Fungsi keputusan klasifikasi $\text{sign}(f(x))$:</p> $f(x) = w \cdot x + b \text{ atau}$ $f(x) = \sum_{i=1}^m a_i \cdot y_i K(x, x_i) + b$ <p>5. Persamaan kernel:</p> <p>a. <i>Kernel Linear</i></p> $K(x, y) = x \cdot y$ <p>Keterangan:</p> <ul style="list-style-type: none"> - y_i = banyak data y sejumlah i - x = vector data (input data x) - w = bobot vector - b = bias - a_i = nilai bobot setiap titik data - $K(x, x_i)$ = Fungsi kernel

2.2.3 *Text Preprocessing*

Text Preprocessing merupakan proses awal sebelum melakukan klasifikasi. Proses ini digunakan untuk membersihkan data dari noise agar data siap digunakan pada proses selanjutnya. (Fadhilah Dwi Ananda, Yoga Pristyanto, 2021)

2.2.4 *Evaluasi*

Dalam hal klasifikasi teks terkadang terdapat kesalahan dalam mengklasifikasinya. Oleh sebab itu, perlu diketahui seberapa besar akurasinya untuk dibandingkan dengan penelitian lain. Salah satu cara untuk mengetahui besar akurasi pengklasifikasian teks adalah menggunakan perhitungan akurasi. Akurasi adalah tingkat kedekatan antara nilai prediksi dan nilai aktual. (Ferdinand Alvianda, 2018)

$$\text{Akurasi} = (TP + TN) / (TP + TN + FP + FN)$$

Keterangan:

TP = True Positive

TN = True Negative

FP = False Positive

FN = False Negative

2.2.5 *Word Cloud*

Word Cloud adalah visualisasi gambar yang menunjukkan daftar kata-kata yang digunakan dalam sebuah teks, umumnya semakin banyak kata yang digunakan semakin besar ukuran kata tersebut dalam gambar. (Fani Widodo, 2021)