

BAB II
LANDASAN TEORI

2.1 Tinjauan Pustaka

Peneliti menampilkan Tabel *review* jurnal-jurnal yang menjadi referensi dalam penelitian ini, berikut Tabel ditampilkan:

Tabel 2.1 Studi Pustaka

No	Peneliti (Tahun)	Judul	Algoritma	Hasil
1.	Wanda Athira Luqyana, Imam Cholissodin, Rizal Setya Perdana (2018)	Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine	<i>Support Vector Machine</i>	Metode SVM saat menganalisis emosi cyberbullying di kolom komentar instagram dapat membagi menjadi kelas sentimen positif dan sentimen negatif cyberbullying. Berdasarkan hasil tes yang dijalankan tingkat akurasi tertinggi sebesar 90%, precision sebesar 94,44% dan recall 85% sebesar dengan komposisi data latih 50% dan data uji 50%.

2.	Siti Nurhasanah Nugraha, Tri Rivanie, Sri Rahayu, Windu Gata, Rangga Pebrianto (2020)	Sentimen Analisis Penerapan <i>Social Distancing</i> Menggunakan <i>Feature Selection</i> Pada Algoritma <i>Support Vector Machine</i>	<i>Support Vector Machine</i>	Dari penelitian pada jurnal kali ini diketahui bahwa hasil cross validation algoritma <i>SVM</i> dengan penambahan <i>feature selection</i> memiliki nilai akurasi mencapai 70,33%. Dari nilai tersebut diketahui bahwa peningkatan dengan algoritma yang dipakai nilai akurasi sebesar 3,33%.
3.	Syifa Rahmawati Hakim, M. Alfa Rizki, Noval Irgi Zekha F, Nurhidayatul Fitri, Yolanda Rizkie A, Rani Nooraeni (2020)	Analisis Sentimen Pengguna Instagram Terhadap Kebijakan Kemdikbud Mengenai Bantuan Kuota Internet dengan Metode <i>Support Vector Machine</i> (SVM)	<i>Support Vector Machine</i>	Dari data yang telah crawling sebanyak 1483 komentar instagram terhadap kebijakan Kemdikbud mengenai dukungan kuota internet selama pandemi COVID 19 diidentifikasi sebagai 70,67% sentimen positif dan 29,33% sentimen negatif. Selanjutnya diklasifikasikan berdasarkan data awal (Total data), data pelatihan, data

				<p>pengujian Model SVM Radial (RBF) Menghasilkan Rata-rata Persentase komentar berupa sentimen positif hingga 61,5% Ini menunjukkan bahwa Jawaban yang bagus dan terima kasih pengguna Instagram tentang kebijakan Kementerian Pendidikan dan Kebudayaan Tentang pemberian dukungan kuota internet selama periode tersebut Pandemi covid19. 38,5% lainnya Itu adalah komentar dalam bentuk sentimen negatif. menampilkan keluhan kepada pengguna instagram terhadap kebijakan Kementerian Pendidikan dan Kebudayaan. Hal ini dibuktikan dengan tingkat akurasi</p>
--	--	--	--	--

				sebesar 79.67%, sensitivitas sebesar 78.89%, dan spesifisitas sebesar 81.82%.
4.	Styawati, Nirwana Hendrastuty, Auliya Rahman Isnain, Ari Yanti Rahmadhani (2021)	Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine	<i>Support Vector Machine</i>	Diskusi tentang program kartu pra-kerja di twitter sangat menarik untuk dianalisa tingkat sentimen masyarakat terhadap program tersebut adalah upaya rintisan pemerintah dalam mengatasi pengangguran. Data yang diperoleh di media sosial twitter Selanjutnya diklasifikasikan menggunakan Support Vector Machine (SVM). Sedangkan dalam mengukur kinerja klasifikasi SVM menggunakan metode Confusion Matrix. Sebuah perbandingan dibuat dalam penelitian ini Dua kernel, RBF dan

				<p>linier. Hasil evaluasi dieksekusi dengan nilai akurasi linear kernel 98,67%, akurasi 98%, recall 99%, F1 Score 98%, Akurasi kernel RBF 98,34%, akurasi 97%, callback 98%, F1 Dengan peringkat 98%, opini publik Pengguna Twitter tentang program kartu pra-pembelian di masa mendatang Pandemi adalah 98,34%, yang cukup netral. berdasarkan Hasil evaluasi yang dilakukan pada nilai akurasi linier kernel Kernel RBF menghasilkan nilai akurasi 98,67% Anda mendapatkan akurasi 98,34%. Artinya, dari perspektif akurasi kernel Penyelesaian lebih akurat daripada kernel RBF.</p>
--	--	--	--	---

5.	Primandani Arsi, Retno Waluyo (2019)	Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma <i>Support Vector Machine</i> (SVM)	<i>Support Vector Machine</i>	Salah satu topik yang paling ramai dibahas di Twitter adalah tentang pemindahan ibu kota Indonesia. Namun dibalik hal tersebut terdapat kontroversi dari pihak-pihak yang merasakan pro dan kontra masing-masing dengan perspektif yang berbeda. Hal ini menimbulkan fenomena kontroversial, terutama di twitter. Ini sebenarnya mewakili kepedulian kolektif untuk wacana publik. Dalam hal ini Analisis sentimen digunakan untuk mengekstraksi, memahami, dan memproses data dalam bentuk teks yang tidak terstruktur secara otomatis untuk mengambil informasi emosional yang terkandung dalam frasa atau opini. Ada
----	--	---	---------------------------------------	---

				<p>beberapa teknik yang banyak digunakan untuk menerapkan analisis sentimen menggunakan teknik pembelajaran mesin. Penelitian ini mengusulkan untuk menerapkan metode <i>Support Vector Machine</i> (SVM) pada tweet tentang perjalanan di ibukota Indonesia untuk mengklasifikasikan kelas sentimen di media sosial Twitter. Klasifikasi teknis didasarkan pada klasifikasi menjadi dua kelas, positif dan negatif. Berdasarkan hasil pengujian yang dilakukan pada tema transisi dari social media Twitter, Akurasi diperoleh dengan SVM = 96.68, precision=95.82% dan recall=94.04%</p>
--	--	--	--	--

6.	Dedi Darwis, Eka Shintya Pratiwi, A. Ferico Octaviansyah Pasaribu (2020)	Penerapan Algoritma SVM untuk Analisis Sentimen pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia	<i>Support Vector Machine</i>	Metode klasifikasi yang digunakan dalam penelitian ini merupakan <i>Support Vector Machine</i> (SVM) dan ekstraksi fitur menggunakan TF-IDF. Hasil preprocessing guna menghitung nilai kemunculan kata dari label yang sesuai lalu menghasilkan sentimen yang positif, negatif dan netral. Berdasarkan hasil tes Penerapan metode SVM yang dihasilkan memberikan nilai akurasi sebesar 82%. Hasilkan sentimen label netral 25% label negatif lebih besar dengan jumlah 77%, dan label positif 8%
7.	Auliya Rahman Isnain, Adam Indra Sakti, Debby Alita, Nurman Satya Marga (2021)	Sentimen Analisis Publik Terhadap Kebijakan <i>Lockdown</i> Pemerintah Jakarta	<i>Support Vector Machine</i>	Dari penelitian kali ini diketahui bahwa metode <i>Support Vector Machine</i> memiliki nilai akurasi sebesar 74%. Menggunakan fitur tf-

		Menggunakan Algoritma <i>Support Vector Machine</i>		idf dengan rincian Accuracy (74%), Precision (75%), dan Recall (92%)
8.	Muhammad Fikry, Yusra (2018)	Ekstrovert atau Introvert: Klasifikasi Kepribadian Pengguna Twitter dengan Menggunakan Metode <i>Support Vector Machine</i>	<i>Support Vector Machine</i>	Untuk mengklasifikasikan kepribadian, algoritma <i>SVM</i> cukup berhasil dalam menganalisa kelas introvert dan ekstrovert. Dapat dikatakan berhasil karena hasil dari akurasi modelnya menembus angka akurasi sebesar 88,89%.
9.	Ahmad Roihan, Po Abas Sunarya, Ageng Setiani Rafika(2020)	Pemanfaatan Machine Learning dalam Berbagai Bidang		Evaluasi dari berbagai disiplin ilmu disajikan dalam bentuk pemecahan masalah termasuk pembelajaran penguatan. Tinjauan terbatas pada beberapa bidang, dan hasilnya menunjukkan bahwa obat-obatan atau obat-obatan saat ini paling dominan, bersama dengan

				banyak bidang lain seperti industri, teknologi, transportasi dan hal ini tidak menutup kemungkinan untuk menganalisa bidang lain yang belum pernah terjamah oleh pendekatan machine learning.
--	--	--	--	---

2.2 Klasifikasi

Komputasi merupakan salah satu hal terpenting dalam teknologi informasi. Berbagai jenis data yang diolah secara efektif dan efisien memberikan informasi yang akurat tentang. Seiring dengan kemajuan teknologi, berbagai metode pengolahan data juga telah dikembangkan. Ini adalah untuk memenuhi kebutuhan metode yang digunakan menurut beserta jenis data yang sedang diproses. Salah satu metode yang digunakan dalam pengolahan data adalah klasifikasi. Klasifikasi adalah metode pengelompokan objek berdasarkan karakteristik objek klasifikasi selain itu klasifikasi juga dapat dikatakan salah satu tugas utama dalam data mining. Pengklasifikasi dibuat dari kumpulan data yang dicari berdasarkan kategori tertentu dengan kelas yang telah ditentukan sebelumnya. Klasifikasi dapat dilakukan secara manual atau dengan bantuan teknologi dengan berbagai cara. Klasifikasi yang dieksekusi secara manual adalah klasifikasi yang dilakukan oleh manusia tanpa menggunakan algoritma cerdas komputer. Klasifikasi dilakukan dengan bantuan teknologi, tetapi ada beberapa algoritma termasuk pendekatan yang salah satunya adalah *machine learning*.

2.3 Analisis Sentimen

Analisis Sentimen adalah jenis klasifikasi teks yang mengkategorikan bagan (memisahkan) teks berdasarkan penyesuaian sentimen opini yang terdapat di

dalamnya. Hal ini merupakan bagian dari *Natural Language Processing*, NLP sendiri adalah salah satu bidang ilmu komputer yang tujuannya untuk menciptakan kecerdasan buatan guna mengetahui interaksi bahasa manusia dan komputer. Analisis sentimen pada prosesnya dapat mendeteksi kata-kata secara spesifik yang nantinya dapat memisahkan serta mengklasifikasi teks ke dalam kelas-kelas tertentu dan kelas yang biasa dipakai adalah opini positif, negatif dan netral. Hal ini juga disebut dalam penggalan opini yang fungsinya memperoleh sikap dan pendapat pembicara. Untuk menelaah hal ini, pendapat dikumpulkan bersumber dari pengguna media sosial. Media sosial beroperasi sebagai instrumen yang mana pemakai (user) bisa memposting banyak pendapat setiap hari dan portal ini yang nantinya digunakan untuk klasifikasikan dari teks yang sudah tersedia (posting). Analisis sentimen memiliki peran penting guna catatan dalam kinerja, mengukur kualitas dan menjadi peran besar dalam evaluasi yang cukup akurat (Devika, Sunirta, Amal Ganesh, 2016). Hal ini juga dikarenakan teks yang akan dianalisa nantinya merupakan hubungan timbal balik langsung dari masyarakat yang mana opini yang ditulis dalam keadaan dan luapan emosi yang beragam terhadap objek tersebut. Tujuan akhir dari sentimen analisis adalah menghasilkan informasi sentimen yang sangat halus yang dapat bermanfaat dalam mengevaluasi hasil dari kinerja yang telah dilakukan.

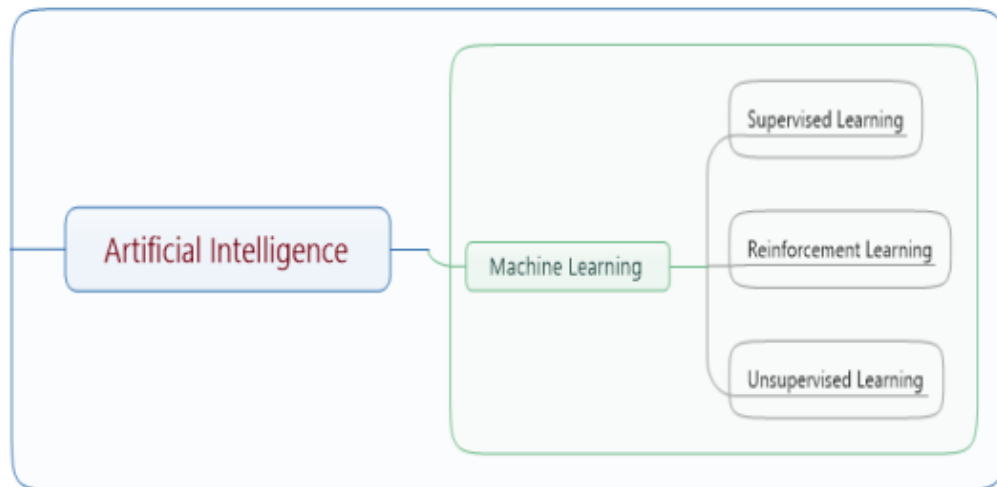
2.4 Pemerintahan Daerah

Sistem pemerintahan adalah sistem hukum ketatanegaraan baik yang berbentuk monarki maupun republik dalam hubungannya dengan hubungan antara pemerintah dengan badan-badan yang mewakili rakyat. Untuk di Indonesia, mengikuti prinsip "check and balance" dan menganut pemisahan kekuasaan, sistem presidensial dengan departemen administrasi, legislatif dan yudikatif. Ketentuan ini termasuk dalam UUD, namun masih perlu perbaikan, terutama regulasi yang lebih jelas. Pemisahan kekuasaan dan pembatasan wewenang di antara ketiga lembaga negara. Di bidang pemerintahan daerah itu sendiri, negara biasanya dijalankan oleh daerah otonom dan gubernur yang menjalankan hak-hak daerah otonom, dan dalam menjalankan kebijaksanaannya, wewenang untuk mengelola dan menyelidiki potensi kekayaan alam itu juga. Sumber daya rakyat untuk mencapai tujuan dan

sasaran otonomi daerah dengan baik. Komitmen, dukungan dan partisipasi masyarakat dalam penciptaan ide dan kreativitas berupa gagasan untuk kemajuan daerah itu sendiri. Bagi pemerintah daerah sendiri perlu memberikan hak dan wewenang tersebut serta melaksanakannya dengan penuh tanggung jawab. Otonomi. Pengelolaan dan pemanfaatan sumber daya alam harus dilakukan dengan hati-hati jika konsekuensinya adalah pendapatan tambahan bagi keuangan daerah, yang untuk kepentingan masyarakat luas. Demikian pula halnya dengan pengelolaan dan partisipasi masyarakat dalam pelaksanaan otonomi daerah. Dalam pelaksanaan otonomi daerah, pelaksanaan demokrasi menuntut kualitas pemimpin daerah yang bertugas menentukan sukses tidaknya penyelenggaraan daerah dan harus digugurkan dalam pemilihan kepala daerah. Otonomi. Oleh karena itu, sistem otonomi dapat dikatakan berhasil di daerah, tergantung pada kemampuan sumber daya manusia dalam mengelola potensi sumber daya alam yang dimiliki oleh daerah otonom.

2.5 Pendekatan Machine Learning dan Algoritma SVM

Kecerdasan buatan merupakan salah satu bidang ilmu komputer yang bertujuan untuk mengembangkan perangkat lunak dan perangkat keras yang dapat dianggap mampu atau mendekati cara kerja otak manusia (Sunarya, Santoso & Sentanu, 2015). Machine learning merupakan salah satu cabang dari kecerdasan buatan yang telah banyak diteliti dan digunakan untuk memecahkan berbagai masalah. Seiring berjalannya waktu, mesin pintar atau Artificial intelligence secara bertahap akan menggantikan dan meningkatkan keterampilan manusia di berbagai bidang (Nayak & Dutta, 2017). Secara umum, pembelajaran mesin dapat didefinisikan sebagai aplikasi komputer dan algoritma matematika yang diadopsi dengan mempelajari dan memprediksi masa depan yang berasal dari data (Goldberg & Holland, 1988). Bidang pembelajaran mesin menjawab pertanyaan tentang bagaimana menulis program komputer untuk perbaikan otomatis berdasarkan pengalaman (Mitchell, 1997). Menurut sebuah penelitian baru-baru ini, pembelajaran mesin terbagi dalam tiga kategori: pembelajaran yang *Supervised Learning*, *Unsupervised Learning*, *Reinforcement Learning*. Bagian metode dalam Machine Learning dapat dilihat pada bagan dibawah ini:



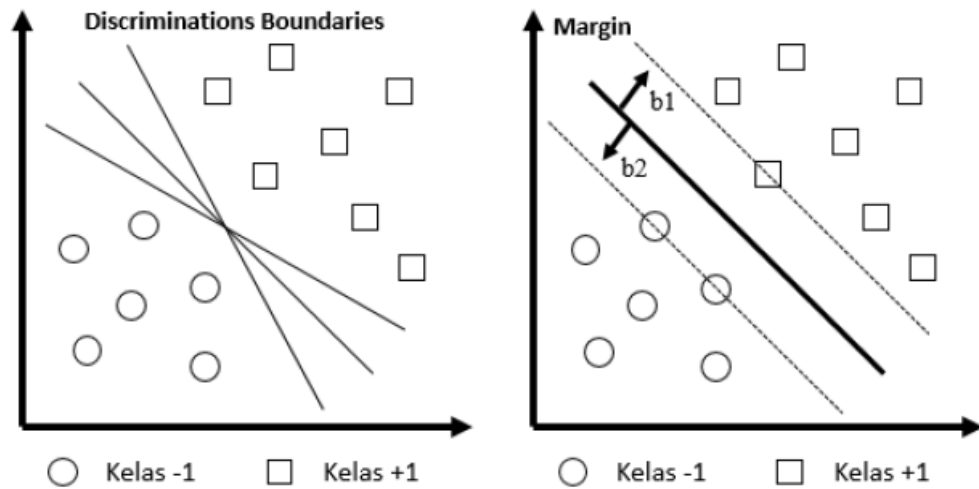
Gambar 2.1 Penggolongan AI

Teknik yang digunakan dalam pembelajaran terawasi adalah metode klasifikasi yang sepenuhnya melabeli catatan untuk mengklasifikasikan kelas yang tidak diketahui. Teknik unsupervised learning sering disebut sebagai cluster karena tidak memerlukan pelabelan dataset dan tidak mengidentifikasi contoh kelas dengan hasil yang ditentukan (Thupae, Isong, Gasela & Abu Mahfouz, 2018). Pembelajaran penguatan biasanya bergerak antara pembelajaran yang diawasi dan tidak diawasi tetapi dengan teknik ini, konsep menetapkan tujuan tanpa secara eksplisit memberitahu komputer ketika tujuan tercapai. Bekerja dalam lingkungan dinamis yang perlu dicapai. Sebagai peneliti, penelitian ini hanya berfokus pada metode pembelajaran terawasi dengan algoritma SVM, SVM sendiri akan dijelaskan sebagai berikut:

2.5.1 Support Vector Machine

SVM merupakan salah satu algoritma dalam memprediksi klasifikasi dan regresi yang paling populer saat ini. Dukungan mesin vektor menggunakan konsep bidang keputusan dalam menentukan batas-batas keputusan. Tingkat keputusan adalah tingkat di mana Anda memisahkan set objek dengan keanggotaan kelas yang berbeda. *SVM* memiliki kemampuan untuk memecahkan masalah linier dan non-linier. Pendekatan kernel yang dikembangkan dengan *SVM* juga dapat digunakan untuk berbagai kelas. Secara teoritis, *SVM* dirancang untuk dua kelas klasifikasi

yang bertujuan menemukan hyperplane yang paling optimal. *Hyperplane* adalah fungsi pemisahan antara dua kelas ruang input (Vapnik, 1992).



Gambar 2.2 Hyperplane

Pada metode SVM pemisahan adalah fungsi linear. Semua data pelatihan diwakili oleh (x_i, y_i) . Setiap data latih dinyatakan oleh (x_i, y_i) dengan $i = 1, 2, \dots, N$, dan $x_i = \{x_{i1}, x_{i2}, \dots, x_{iq}\}$ T merupakan atribut (fitur) set untuk data latih kelas ke- i . $-1, +1$ menyatakan label kelas. Umumnya, hyperplane direpresentasikan sebagai $w \cdot x + b = 0$. Untuk masalah klasifikasi yang umum, sampel data tidak terpisah secara linear, sehingga hasil yang diperoleh saat menggunakan mesin vektor pendukung linear tidak optimal dan hasil klasifikasi buruk. Mesin vektor pendukung linier dapat diganti dengan mesin vektor pendukung non-linier menggunakan metode kernel. Fungsi yang dapat digunakan pada umumnya terdapat empat jenis, yaitu:

a. Kernel Linear

$$K(x, x_k) = x_k^t x$$

b. Kernel Polynomial

$$K(x, x_k) = (x_k^t x + 1)^d$$

c. Kernel Gaussian (radial basis function, RBF)

$$K(x, x_k) = \exp \left\{ -\frac{\|x - x_k\|^2}{2\sigma^2} \right\}$$

d. Kernel Sigmoid

$$K(x, x_k) = \tanh[k x_k^t x + 0]$$

Penjelasan diatas merupakan jenis binary class dan dalam perkembangnya metode SVM tidak hanya dapat mengklasifikasikan dalam dua kelas (binary class SVM) akan tetapi metode SVM dapat mengklasifikasikan lebih dari dua kelas (multi class SVM). Beberapa metode yang menggunakan cara klasifikasi ini adalah:

- One Against-All

Metode ini berkonsep kepada membandingkan salah satu kelas dengan seluruh kelas. Dalam mengklasifikasikan data kedalam kelas k kelas, sejumlah model SVM biner harus dibangun sebelumnya. Contohnya adalah, untuk mengklasifikasikan data ke dalam 4 kelas, kita harus membangun 4 kelas biner terlebih dahulu. Kelemahan dari metode ini adalah metode ini dapat membentuk area baru yang tidak dapat diklasifikasikan dari kelas sebelumnya.

- One Against-All

Metode ini berkonsep kepada model SVM yang membandingkan kelas dengan kelas lainnya. Dalam menggunakan metode ini, pendekatan yang digunakan adalah: $\frac{k(k-1)}{2}$. Dengan rumus tersebut kita dapat mengklasifikasikan kelas yang ada. Dengan kata lain saat kita ingin menyelesaikan klasifikasi 4 kelas artinya kita harus membangun 6 buah SVM biner. Setiap kelas harus dibandingkan dengan tiga kelas lainnya dan dapat dianalogikan dengan kompetisi penuh dalam pertandingan sepakbola. Dapat disimpulkan kelas yang paling sering menang merupakan kelas keputusan. Adapun jika kelas tersebut memiliki jumlah kemenangan yang salah maka kelas putusan dapat dipilih secara acak dan dapat juga dengan memberikan pembobotan nilai secara tertentu.

- Decision Directed Acyclic Graph

Dalam pengklasifikasiannya metode ini mirip dengan metode sebelumnya, hanya saja metode DDAG dalam hal ini menggunakan sistem gugur. Kelas yang mengalami kekalahan tidak akan dibandingkan dengan tahapan selanjutnya. Dengan catatan kelas yang sudah menang berkali-kali dapat saja kalah pada tahap terakhir. Dan dalam penerapannya metode ini sangat sensitif terhadap urutan.

- Adaptive Direct Acyclic Graph

Didesain untuk mengatasi kekurangan DDAG yang cukup sensitif terhadap urutan klasifikasi antar kelas. Metode ini dapat dianalogikan dengan kelas keputusan harus bertanding penuh mulai dari babak penyisihan, semifinal sampai final dengan kata lain tidak ada kelas yang tiba tiba muncul di babak akhir dengan sekali kemenangan. Dengan model seperti ini ADAG lebih tahan terhadap urutan.

Seperti yang dijelaskan (Aji Prasetya Wibawa, 2018) SVM memiliki kekurangan dan kelebihan, kelebihan dan kekurangan tersebut akan dijelaskan sebagai berikut: Sebagai salah satu algoritma SVM kelebihan yaitu:

- Mengklasifikasikan suatu pattern, yang tidak termasuk data yang dipakai dalam fase pembelajaran metode itu
- Dapat diimplementasikan dengan mudah kedalam satu model

Selain itu, SVM juga memiliki kekurangan yaitu:

- Sulit atau kurang cocok dipakai dalam masalah berskala besar. Skala besar dalam hal ini dimaksudkan dengan jumlah sampel (data) yang diolah
- Jenis kernel SVM berpengaruh pada akurasi sistem

2.6 Instagram

Instagram merupakan media sosial yang pada dasarnya berfungsi untuk berbagi foto maupun video pada sesama pengguna. Sejak September 2017 pengguna aktif instagram telah tercatat kurang lebih 800 juta orang (Yusuf, 2017). Pengguna instagram tidak hanya terdiri dari kalangan tertentu saja, pengguna instagram menjangkau banyak user termasuk hampir semua kalangan masyarakat Provinsi

Lampung. Pemanfaatan yang beragam dari pengguna Instagram dibuat sebagai akun pribadi, baik untuk orang biasa maupun artis hingga menjadi sarana bisnis perseorangan atau juga bisa menjadi akun official dari suatu instansi atau organisasi tertentu. Provinsi Lampung termasuk instansi yang memiliki akun official yang digunakan untuk memberitahu masyarakat kegiatan apa saja yang sedang dijalankan oleh Gubernur Provinsi Lampung. Dalam instagram sendiri memiliki fitur komentar hal ini tentu dipakai beberapa masyarakat untuk menyampaikan persepsinya terhadap Provinsi Lampung, termasuk menyampaikan keluhan kesah sampai memuji kebijakan yang sedang dijalankan oleh Gubernur Arinal Djunaidi.