

## BAB II TINJAUAN PUSTAKA

### 2.1. Penelitian Terdahulu

Penelitian yang memiliki tujuan untuk mengelompokkan, mengolah, menyelesaikan beraneka macam data serta kasus dalam berbagai bidang di Indonesia maupun luar negeri. Pada penelitian yang telah dilakukan sebelumnya akan membantu untuk penelitian selanjutnya agar dapat menghasilkan penelitian yang lebih baik lagi.

Dalam penelitian yang dilakukan oleh Tri Mueri Sandes yang berjudul “Klasifikasi *Tweet* Pelecehan Online Pada *Twitter* Dengan Menggunakan Metode *Naïve Bayes*” pada tahun 2018. Tentang masalah kejahatan siber seperti *cyber harassment* atau pelecehan online yang terjadi pada pengguna aplikasi *Twitter* menggunakan algoritma *Naïve Bayes* serta hasil akurasi tertinggi 85% dengan perbandingan data yaitu 80 banding 20 untuk data latih dan data uji. Dengan demikian penerapan metode *Naïve Bayes* dapat diimplementasikan untuk kasus pelecehan online pada *Twitter* [15].

Penelitian lain yaitu dilakukan oleh Faisal Aji Prayoga, Aryo Pinandito, dan Rizal Setya Perdana yang berjudul “Rancang Bangun Aplikasi Deteksi Spam *Twitter* menggunakan Metode *Naive Bayes* dan *KNN* pada Perangkat Bergerak *Android*” pada tahun 2018. Penelitian ini merancang kerangka aplikasi *Android* yang memungkinkan pengembang untuk membangun penerapan klasifikasi-klasifikasi spam untuk konten *Twitter* sebagai perpustakaan aplikasi. Metode klasifikasi yaitu *Naive Bayes* dan *K-Nearest Neighbor* dapat mendeteksi kandungan spam dan ham masing masing dengan akurasi 82% dan 71% [20].

Penelitian berikutnya yaitu dilakukan oleh Amer Saeed Ali Al-Katheri & Maheyzah Md Siraj yang berjudul “*Classification of Sexual Harassment on Facebook Using Term Weigthing Schemes*” pada tahun 2018. Penelitian ini tentang pelecehan seksual yang terjadi pada media sosial *Facebook*. Metode yang digunakan *Support Vector Machine* dengan tiga pembobotan yaitu *Entropy*, *TFIDF*, dan *Modified TFIDF*. Penelitian ini menggunakan metode *SVM* untuk mengklasifikasi dua dataset berdasarkan akurasi dan presisi serta hasil percobaan menunjukkan bahwa *Modified TFIDF* lebih tinggi dibandingkan *Entropy* dan *TFIDF* [16].

Dalam penelitian yang dilakukan oleh Irwansyah Saputra dan Didi Rosiyadi yang berjudul “Perbandingan Kinerja Algoritma *K-Nearest Neighbor*, *Naïve Bayes Classifier* dan *Support Vector Machine* dalam Klasifikasi Tingkah Laku Bully pada

Aplikasi Whatsapp” pada tahun 2019. Tentang masalah kejahatan siber seperti *cyberbullying* yang terjadi pada pengguna aplikasi *Whatsapp* menggunakan algoritma *K-Nearest Neighbor*, *Naïve Bayes Classifier* dan *Support Vector Machine* serta hasil menunjukkan bahwa algoritma SVM lebih baik dalam menyelesaikan kasus ini dengan akurasi sebesar 81.58% [21].

Penelitian berikutnya yang dilakukan oleh Risa Wati dan berjudul “Penerapan Algoritma *Naïve Bayes* dan *Particle Swarm Optimization* Untuk Klasifikasi Berita *Hoax* Pada Media Sosial” pada tahun 2020. Berisi tentang dampak negative yang dapat disebabkan oleh pengguna media social salah satunya menyebarkan berita bohong atau *hoax*. Menggunakan metode *Naïve Bayes* dan *Particle Swarm Optimization* yang di padukan, karena hanya menggunakan *Naïve Bayes* masih kurang maksimal. Menggunakan *Naïve Bayes* dan *Particle Swarm Optimization* menghasilkan nilai akurasi sebesar 85.19% [22]

Penelitian yang dilakukan oleh Hermanto, Ali Mustopa Antonius dan Yadi Kuntoro dengan judul “Algoritma Klasifikasi *Naïve Bayes* dan *Support Vector Machine* dalam Layanan Komplain Mahasiswa” pada tahun 2020. Penelitian ini tentang mengklasifikasi antara pernyataan merupakan *complain* atau *not complain* dengan dua algoritma yaitu *Naïve Bayes* dan *SVM*. Hasil perbandingan algoritma tersebut *SVM* lebih tinggi dengan akurasi 84.45% [17].

Penelitian yang dilakukan oleh Enda Esyudha Pratama dan Bambang Riyanto Trilaksono dengan judul “Klasifikasi Topik Keluhan Pelanggan Berdasarkan *Tweet* Menggunakan Penggabungan Feature Hasil Ekstrasi pada Metode *Support Vector Machine* (SVM)” pada tahun 2015. Penelitian ini tentang mengklasifikasi keluhan para pelanggan melalui tweet pada Twitter. ini dapat dibuktikan dari akurasi 82.50% klasifikasi bentuk keluhan, dan 86.67% klasifikasi jenis keluhan [23].

Penelitian yang dilakukan oleh Bening Herwijayanti, Dian Eka Ratnawati dan Lailil Muflikhah dengan judul “Klasifikasi Berita Online dengan menggunakan Pembobotan TF-IDF dan Cosine Similarity” pada tahun 2018. Pada penelitian ini tentang kasus kesalahan manusia terhadap pengkategorian terhadap berita *online* dan hanya menggunakan pembobotan *tf-idf* dan *cosine similarity*. Memiliki tujuan untuk mempermudah human eror serta kejadian kesalahan pengkategorian dan klasifikasi mampu mengelompokkan berita dengan tingkat akurasi 91.25% [24].

Penelitian yang dilakukan oleh Tifani Wulandari dengan judul “Klasifikasi Jenis Emosi dari Tweet Berbahasa Indonesia Menggunakan Support Vector Machine” pada tahun 2018. Penelitian ini tentang mengklasifikasi jenis emosi para pengguna Twitter yang yang membuat *tweet*. Model pembelajaran menggunakan SVM, hasil akurasi sebesar 85% dengan treshold 30 dan menggunakan  $C=2$  dan  $\text{Gamma}=0.1$  [25].

Tabel 2.1 Penelitian terdahulu

No.	Judul penelitian	<i>Comparing</i>	<i>Contrasting</i>	<i>Criticize</i>	<i>Synthesize</i>	<i>Summarize</i>
1.	KLASIFIKASI <i>TWEET</i> PELECEHAN ONLINE PADA <i>TWITTER</i> DENGAN MENGUNAKA N METODE <i>NAÏVE BAYES</i>	Kasus pelecehan online atau <i>cyber</i> <i>harassment</i> masih banyak terjadi di twitter menggunakan metode <i>Naïve</i> <i>Bayes</i> .	Pada penelitian ini hanya menggunakan satu metode yaitu <i>naïve bayes</i> serta membandingkan lima kelas seperti <i>racist,</i> <i>embarrass,</i> <i>physically</i> <i>threatened,</i> <i>sexually harassed</i> atau <i>neutral</i> .	Penelitian ini hanya menggunakan satu metode maka belum mengetahui nilai akurasi menggunakan metode lainnya. Sepertinya penulis belum menjadikan penelitian ini sebagai paper dalam jurnal.	<i>Naïve Bayes</i> <i>Classifier</i>	Berdasarkan hasil penelitian dengan tahapan text preprocessing tanpa stemming diperoleh hasil akurasi tertinggi 85% pada perbandingan data 80:20 untuk data latih dan data uji. Dengan demikian penerapan metode <i>Naïve Bayes</i> dapat diimplementasikan untuk kasus pelecehan online pada Twitter.
2.	Rancang Bangun Aplikasi Deteksi Spam <i>Twitter</i> menggunakan	Penelitian ini merancang kerangka aplikasi Android	Penelitian ini mengenai masalah spam yang terjadi di	Dataset yang digunakan masih relative sedikit yaitu	<i>K-Nearest</i> <i>Neighbor</i> dan <i>Naïve Bayes</i>	Metode klasifikasi <i>Naive Bayes</i> dan <i>K-Nearest Neighbor</i> dapat mendeteksi

	Metode <i>Naive Bayes</i> dan <i>KNN</i> pada Perangkat Bergerak <i>Android</i>	yang memungkinkan pengembang untuk membangun penerapan klasifikasi spam untuk konten Twitter menggunakan Metode <i>Naive Bayes</i> dan <i>KNN</i> .	Twitter dan membandingkan dua metode <i>Naive Bayes</i> dan <i>KNN</i> .	antara 10-100 <i>tweet</i> .		kandungan spam dan ham masing masing dengan akurasi 82% dan 71%.
3.	<i>Classification of Sexual Harassment on Facebook Using Term Weigthing Schemes</i>	Penelitian ini tentang pelecehan seksual yang terjadi pada media sosial <i>Facebook</i> .	Perbedaan yaitu berupa pada metode yang hanya menggunakan satu metode dan sosial media <i>Facebook</i> .	Pada abstrak maupun di kesimpulan, penulis di penelitian ini tidak menjelaskan hasil akurasi berupa presentase.	Algoritma <i>Support Vector Machine</i> , dengan tiga pembobotan term <i>Entropy</i> , <i>TFIDF</i> , <i>Modified TFIDF</i> .	Penelitian ini menggunakan metode SVM untuk mengklasifikasi dua dataset berdasarkan akurasi dan presisi serta hasil percobaan menunjukkan bahwa <i>Modified TFIDF</i> lebih tinggi dibandingkan

						<i>Entropy</i> dan <i>TFIDF</i> . Hasil akurasi dengan metode SVM dan <i>Modified TFIDF</i> yaitu 30 fitur 78%, pada 60 dan 90 fitur 96.50%.
4.	Perbandingan Kinerja Algoritma <i>K-Nearest Neighbor</i> , <i>Naïve Bayes Classifier</i> dan <i>Support Vector Machine</i> dalam Klasifikasi Tingkah Laku Bully pada Aplikasi Whatsapp	Kejahatan siber seperti <i>Cyberbullying</i> yang terjadi pada pengguna aplikasi <i>Whatsapp</i> menggunakan Algoritma <i>K-Nearest Neighbor</i> , <i>Naïve Bayes Classifier</i> dan <i>Support Vector Machine</i> .	Penelitian ini tentang kasus <i>Cyberbullying</i> pada aplikasi <i>whatsapp</i> dan membandingkan tiga metode.	Penulis tidak menulis masing-masing perbedaan algoritma yang digunakan pada penelitian ini secara teori nya.	<i>K-Nearest Neighbor</i> , <i>Naïve Bayes Classifier</i> dan <i>Support Vector Machine</i>	Hasil menunjukkan bahwa algoritma SVM lebih baik dalam menyelesaikan kasus ini dengan akurasi sebesar 81.58%.
5.	Penerapan Algoritma <i>Naïve Bayes</i> dan <i>Particle</i>	Dampak negative yang dapat disebabkan	Penelitian ini tentang kasus berita <i>hoax</i> pada	Tidak menjelaskan bagian landasan	Algoritma <i>Naïve Bayes</i> dan <i>Particle</i>	Menggunakan <i>Naïve Bayes</i> dan <i>Particle Swarm</i>

	<i>Swarm Optimization</i> Untuk Klasifikasi Berita <i>Hoax</i> Pada Media Sosial	oleh pengguna media sosial salah satu nya menyebarkan berita bohong atau <i>hoax</i> .	berbagai media sosial.	teori mengenai metode yang digunakan.	<i>Swarm Optimization.</i>	<i>Optimization</i> menghasilkan nilai akurasi sebesar 85.19%.
6.	ALGORITMA KLASIFIKASI <i>NAIVE BAYES</i> DAN <i>SUPPORT VECTOR MACHINE</i> DALAM LAYANAN KOMPLAIN MAHASISWA	Penelitian ini tentang layanan untuk mahasiswa yang berhak untuk <i>complain</i> maka dibutuhkan sesuatu untuk klasifikasi kasus ini menggunakan algoritma <i>NAIVE BAYES</i> dan <i>SUPPORT VECTOR MACHINE</i> .	Perbedaannya ialah pada kasus permasalahan karena dibidang Pendidikan dan terutama subjek nya adalah mahasiswa.	Masih banyak tanda baca yang kurang rapi dan beberapa kesalahan pada pengejaan kata di penulisan.	Algoritma <i>NAIVE BAYES</i> dan <i>SUPPORT VECTOR MACHINE</i> .	Diperoleh algoritma Support vector Machine memiliki nilai akurasi tertinggi dibanding Naïve Bayes. Nilai AUC = 0,922. untuk metode Support vector Machine dengan memiliki akurasi 84.45%, dari algoritma Naïve Bayes memiliki tingkat akurasi sekitar 69.75% dan nilai AUC=0.679.

7.	Klasifikasi Topik Keluhan Pelanggan Berdasarkan <i>Tweet</i> Menggunakan Penggabungan Feature Hasil Ekstrasi pada Metode <i>Support Vector Machine</i> (SVM)	Penelitian ini tentang klasifikasi keluhan pada pelanggan melalui <i>tweet</i> di aplikasi Twitter	Pada penelitian ini hanya menggunakan satu metode yaitu <i>Support Vector Machine</i> serta masalah nya tentang keluhan pelanggan.	Tidak ada diagram alir penelitian.	Metode <i>Support Vector Machine</i> , metode ekstrasi <i>term frequency</i> (TF) dan <i>document frequency</i> (DF)	Hasil penelitian menunjukkan bahwa SVM mampu melakukan klasifikasi dengan baik, hal ini dapat dibuktikan dari akurasi 82.50% klasifikasi bentuk keluhan, dan 86.67% klasifikasi jenis keluhan.
8.	Klasifikasi Berita <i>Online</i> dengan menggunakan Pembobotan <i>TF-IDF</i> dan <i>Cosine Similarity</i>	Penelitian ini tentang klasifikasi suatu berita <i>online</i> menggunakan pembobotan <i>tf-idf</i> dan <i>cosine similarity</i> .	Pada penelitian ini tentang kasus kesalahan manusia terhadap pengkategorian terhadap berita <i>online</i> dan hanya menggunakan pembobotan <i>tf-idf</i> dan <i>cosine similarity</i> .	Pada penelitian hanya sampai tahap <i>text preprocessing</i> dan pembobotan.	Metode pembobotan <i>tf-idf</i> dan <i>cosine similarity</i> .	Memiliki tujuan untuk mempermudah human eror serta kejadian kesalahan pengkategorian dan klasifikasi mampu mengelompokkan berita dengan tingkat akurasi 91.25%.



9.	Klasifikasi Jenis Emosi dari Tweet Berbahasa Indonesia Menggunakan <i>Support Vector Machine</i>	Penelitian ini tentang mengklasifikasi jenis emosi para pengguna Twitter yang membuat <i>tweet</i> .	Perbedaan kasus dengan klasifikasi jenis emosi.	Penelitian ini hanya menggunakan satu metode maka belum mengetahui nilai akurasi menggunakan metode lainnya. Sepertinya penulis belum menjadikan penelitian ini sebagai paper dalam jurnal karena masih sebuah thesis.	Metode <i>SUPPORT VECTOR MACHINE</i>	Model pembelajaran menggunakan SVM, hasil akurasi sebesar 85% dengan treshold 30 dan menggunakan $C=2$ dan $\text{Gamma}=0.1$ .
----	--	--	---	--	--------------------------------------	---

## 2.2. Landasan Teori

Pada penelitian Perbandingan Metode *Naïve Bayes Classifier* dan *Support Vector Machine* Untuk Klasifikasi *Cyber Harassment* Pada *Twitter* mengacu beberapa landasan teori yang terkait. Landasan teori sebagai berikut:

### 2.2.1. *Cyber Harassment*

Undang-Undang Nomor 11 tahun 2008 tentang Informasi dan Transaksi Elektronik sebagaimana telah diubah dengan Undang-Undang Nomor 19 Tahun 2016 (UU ITE) disahkan pada tanggal 21 April 2008 dan menjadi *cyber law* pertama di Indonesia [8]. *Cyber law* atau hukum siber ialah suatu aspek dalam hukum tentang tindakan yang menyimpang dengan menggunakan teknologi dan internet pada dunia maya seperti *cybercrime*. Jenis-jenis kejahatan kategori *cybercrime* yaitu *cyber terorisme*, *cyber pornography* dan *cyber harassment* [4].

*Cyber Harassment* bisa disebut juga dengan pelecehan online dapat berupa mengancam atau melecehkan melalui email, pesan instan, atau memposting informasi secara online. Dalam kasus ini menargetkan orang tertentu, baik dengan secara langsung menghubungi mereka atau dengan menyebarkan informasi pribadi mereka, menyebabkan mereka tertekan, takut, atau marah [5]. Menurut pengertian dalam Kamus Lengkap Bahasa Indonesia, leceh ialah remeh, tidak berharga, hina; lekat dan meleleh pada tempat yang dilalui [26] sedangkan melecehkan artinya memandang rendah (tidak berharga), menghinakan, mengabaikan.

Istilah dari "*Cyber Harassment*" telah digunakan untuk menggambarkan berbagai perilaku yang melibatkan ancaman yang dilakukan berulang-ulang dan *harassment* atau pelecehan dengan penggunaan *e-mail* atau komunikasi secara online lainnya yang membuat orang takut atau khawatir akan keselamatan mereka. *Cyber Harassment* dapat seperti mengancamnya, mengintai atau *stalking* di dunia nyata dan bisa menimbulkan kegelisahan, kesedihan mental, kerusakan fisik, dan bahkan pembunuhan. Macam-macam *cyber harassment* dapat mencakup hal berikut [5]:

1. Mengirimkan *e-mail* atau pesan instan secara berulang yang mungkin dengan tidak langsung dapat mengancam penerima;
2. Memenuhi kotak masuk *e-mail* korban dengan surat yang tidak penting atau diinginkan, atau mengirim berkas korban dengan virus;

3. Menggunakan alamat *e-mail* korban untuk berlangganan ke beberapa layanan online. Mengirim informasi yang keliru dan pesan yang salah ke forum-forum, atau situs tempat kerja korban;
4. Mencuri identitas online seseorang yang digunakan untuk mengirim informasi palsu;
5. Mengirim informasi demografis korban dan mengirim gambar yang berorientasi seksual atau situs pornografi;
6. Mencari dan mengumpulkan berbagai informasi korban yang mungkin telah diposkan oleh situs berita dengan maksud untuk mencari informasi pribadi dan kemudian menggunakan informasi ini untuk melecehkan, mengancam, dan mengintimidasi korban baik secara online maupun di dunia nyata.

Jenis-jenis *cyber harassment* melalui komunikasi online di media sosial sebagai berikut [27]:

1. Pelecehan Seksual (*Sexual Harassment*) ialah ucapan seksual dengan target suatu gender, seperti terhadap wanita. Misalnya, pelaku pelecehan mengomentari tubuh korban secara vulgar atau menyebut hubungan seksual secara agresif.
2. Pelecehan Ras (*Racial Harassment*) ialah merendahkan korban dengan target karakter ras seperti warna kulit, agama, suku, asal negara, budaya dan lain-lain.
3. Pelecehan terhadap Penampilan (*Appearance-related Harassment*) ialah pelecehan yang terkait dengan penampilan menggunakan bahasa yang memalukan mengacu pada penampilan tubuh seseorang. Misalnya dengan ucapan badannya seperti gajah.
4. Pelecehan Intelektual (*Intellectual Harassment*) yaitu menyinggung kekuatan intelektual atau pendapat individu. Misalnya dengan ucapan bodoh, tolol dan sejenisnya.
5. Pelecehan Politik (*Political Harassment*) ialah pelecehan terkait dengan pandangan politik seseorang. Target khas adalah para politisi dan individu yang cenderung membela salah satu politisi tersebut.

### 2.2.2. *Twitter*

*Twitter* ialah jejaring sosial dan mikroblogging dengan fasilitas mengirim teks info atau pesan secara singkat, padat dan realtime pada akun pengguna dengan dalam kalimat maksimum 140 karakter [2]. Namun *Twitter* resmi mengumumkan penambahan batasan kicauan menjadi 280 karakter mulai hari Rabu, 8 November 2017 pada artikel "*Twitter* Resmi Menambah Batasan Kicauan Jadi 280 Karakter" [28]. *Twitter* memiliki keunggulan yang tidak dimiliki oleh situs jejaring lain yaitu persebaran informasi yang lebih cepat, real time dan selalu ada pembaharuan baru. Media jejaring sosial yaitu *twitter* ini didirikan oleh Jack Dorsey, Noah Glass, Biz Stone, dan Evan Williams yang berada di San Fransisco, Amerika Serikat pada tahun 2006 [2]. Sebuah lembaga riset dan media online bernama katadata.co.id tahun 2020 menempatkan *Twitter* pada urutan ke-5 sebagai media sosial yang paling banyak penggunanya di Indonesia setelah *Youtube*, *Whatsapp*, *Facebook* dan *Instagram* [1].

*Twitter* memiliki fasilitas yaitu sebagai pencari (search) untuk mengetahui topik yang sedang ramai dibicarakan oleh pengguna *Twitter*. Ketika ingin mencari tentang *harassment* maka akan muncul berbagai kaitan dengan info atau *tweet* tentang tersebut, baik yang menggunakan tanda pagar (*#Harassment*) maupun yang tidak (*harassment*). Dalam *twitter* tanda pagar/tagar (#) digunakan untuk menjadi kata kunci dalam *tweet* yang dituliskan. Setiap orang yang tertarik dengan tanda pagar tersebut dapat bergabung dengan menuliskan *tweet* yang berhubungan dengan topik di tanda pagar tersebut. Semakin sering tanda pagar tersebut digunakan maka semakin besar peluang topik tersebut menjadi trending topic atau topik yang paling dibahas waktu itu. Untuk mencari *username* atau alamat akun ditandai dengan simbol "@" sehingga contoh nama yang muncul adalah @namaakun.

Dalam *Twitter* terdapat sistem pertemanan (para pengguna akun saling berinteraksi) memiliki istilah yang berbeda seperti sebagai pengikut (*follower*) dan mengikuti (*following*). Jika pengguna masuk ke dalam akun lain dan ingin mengetahui aktifitas dari akun tersebut maka disebut sebagai mengikuti (*following*). Sedangkan jika akun pengguna diikuti oleh akun lain maka akun tersebut disebut pengikut (*follower*).

Pada sistem pertemanan *Twitter* sebagai "mengikuti" (*following*) para pengguna dapat mengikuti sebanyak mungkin akun lain tanpa harus menunggu persetujuan dari pemilik akun [29]. Terkecuali ada pengguna yang melindungi *tweet* nya untuk tidak dibaca oleh banyak orang, maka pengguna bisa melindungi akun miliknya untuk dikunci (*hidden account*) kemudian akun lain

yang ingin mengikuti akun tersebut harus menunggu pemilik akun tersebut menerima sebagai pengikut nya atau tidak.

Meskipun pada hubungan mengikuti akun lain ini tidak serta merta akan mendapatkan timbal balik dengan kata lain tidak semua pemilik akun yang diikuti akan membalas mengikuti namun dengan adanya sistem ini seorang pengguna sudah mendapat manfaat dengan kebebasan mengakses berbagai informasi pada halaman akun-akun ia ikuti, baik akun yang berasal dari dalam maupun luar negeri.

Fasilitas pada *Twitter* yang dapat digunakan untuk menjawab atau mengomentari sebuah *tweet* disebut *reply*. *Reply* digunakan untuk memberi komentar pada satu *tweet* yang dibuat oleh pengguna. Simbol untuk *reply* berbentuk gambar kecil yang terletak pada dibagian bawah *tweet* dan terletak paling kiri. Pengguna *Twitter* dapat secara langsung mengirimkan *tweet* ke pengguna lainnya atau merujuk pada *tweet* akun lain dalam percakapan dengan menggunakan simbol “@”. Simbol “@” dapat digunakan sebagai penanda alamat yang akan dituju, dengan adanya simbol ini menunjukkan bahwa pesan atau *tweet* tersebut ditujukan pada pemilik akun tertentu [29]. Selain itu *Twitter* memiliki fitur yang sejajar dengan *replay* yaitu fitur *Retweet* yang berfungsi sebagai membantu menyebarkan *tweet* tersebut dan fitur *Favorit* berbentuk symbol hati yang menandakan bahwa menyukai *tweet* tersebut.

### 2.2.3. *Machine learning*

*Machine learning* atau pembelajaran mesin ialah metode analisis yang membantu untuk menangani data dalam ukuran besar dengan cara mengembangkan algoritma komputer serta *machine learning* merupakan salah satu cabang dari disiplin ilmu yaitu *Artificial Intellegence (AI)* atau Kecerdasan Buatan tentang pembangunan sistem yang berdasarkan data [30].

*Machine learning* tidak dapat bekerja apabila tidak ada data, maka data sangat penting dan menjadi hal yang pertama kali disiapkan. Data dibagi menjadi dua kelompok yaitu data training (data pembelajaran) dan data testing (data pengujian). *Machine learning* memiliki empat jenis teknik yaitu [18]:

1. *Supervised Learning* ialah melatih model pada data input dan output yang diketahui sehingga dapat memprediksi keluaran masa depan. Teknik ini mamiliki tujuan untuk melakukan klasifikasi serta terdapat label.
2. *Regresi* ialah dapat melakukan prediksi berupa bilangan kontinu.

3. *Semi-supervised Learning* hampir sama dengan *supervised learning*, perbedaan terletak pada pelabelan data yang dapat menggunakan label maupun tidak.
4. *Unsupervised Learning* adalah teknik untuk mencari sifat-sifat (*properties*) data, tidak memerlukan pelabelan.

#### 2.2.4. Data Mining

Data Mining ialah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak dapat diketahui secara manual. Data Mining merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistic, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar [31].

Kata Mining merupakan kiasan dari bahasa Inggris, mine. Jika mine berarti menambang sumber daya yang tersembunyi di dalam tanah, maka Data Mining merupakan penggalian makna yang tersembunyi dari kumpulan data yang sangat besar oleh karena itu Data Mining sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), machine learning, statistik dan basis data.

Data Mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu [32]:

1. *Classification* merupakan suatu teknik dengan melihat pada kelakuan dan atribut dari kelompok yang telah didefinisikan. Teknik ini dapat memberikan klasifikasi pada data baru dengan memanipulasi data yang ada serta telah diklasifikasi dan dengan menggunakan hasilnya untuk memberikan sejumlah aturan. Salah satu contoh yang mudah dan populer adalah *Naives Bayes*, SVM dan *Decision Tree*.
2. *Association* merupakan suatu teknik untuk mengenali kelakuan dari kejadiankejadian khusus atau proses dimana hubungan asosiasi muncul pada setiap kejadian. Salah satu contohnya adalah *Market Basket Analysis*, yaitu salah satu metode asosiasi yang menganalisa kemungkinan pelanggan untuk membeli beberapa item secara bersamaan.
3. *Clustering* dapat digunakan untuk menganalisis pengelompokan berbeda terhadap data, mirip dengan klasifikasi, namun pengelompokan belum didefinisikan sebelum dijalankannya tool data mining. Biasanya menggunakan Metode *Neural Network* atau statistik. *Clustering* membagi

item menjadi kelompok-kelompok berdasarkan yang ditemukan tool data mining.

### **2.2.5. Text Mining**

*Text Mining* ialah suatu teknik yang digunakan untuk menangani masalah klasifikasi, pengelompokan, ekstraksi informasi, dan pencarian informasi [10]. Untuk data *Twitter text mining* dilakukan pada *unstructure text* untuk mengekstraksi teks yang tidak terstruktur menjadi sebuah informasi yang berguna. Proses text mining yang khas meliputi kategorisasi teks, *text clustering*, ekstraksi konsep atau entitas, produksi taksonomi granular, sentiment analysis, penyimpulan dokumen, dan pemodelan relasi entitas. *Text Mining* memiliki sifat hampir sama dengan data mining namun hanya fokus pada teks daripada bentuk data yang lebih terstruktur [11]. Melakukan secara khusus melibatkan penggunaan teknologi *Natural Language Processing* (NLP), yang menerapkan prinsip-prinsip linguistik komputasional untuk menguraikan dan menginterpretasikan set data.

### **2.2.6. Text Preprocessing**

*Text preprocessing* merupakan proses menggali, mengolah dan mengatur informasi dengan cara menganalisis hubungannya dengan aturan-aturan yang ada di data tekstual semi terstruktur atau tidak terstruktur [33]. Untuk lebih efektif dalam proses *text preprocessing*, dilakukan langkah transformasi data ke dalam suatu format yang memudahkan untuk kebutuhan pemakai. Setelah dalam bentuk yang lebih terstruktur dengan adanya proses di atas, data dapat dijadikan sumber data yang dapat diolah lebih lanjut. Tahapan *text preprocessing* sebagai berikut [15]:

1. *Cleaning* yaitu adapun kata atau karakter yang akan dihilangkan berupa simbol, link URL, taggar (#), nama pengguna atau mention (@), *emoticon* dan RT (*Retweet*).
2. *Case Folding* adalah mengubah semua karakter huruf kapital menjadi huruf kecil (lowercase).
3. *Tokenizing* yaitu proses penguraian deskripsi yang semula berupa kalimat-kalimat menjadi kata per kata dan menghilangkan delimiter-delimiter seperti tanda titik (.), koma (,), spasi dan karakter angka yang ada pada kata.
4. *Normalization* merupakan perbaikan dan substitusi kata yang salah eja ataupun disingkat dengan bentuk tertentu. Substitusi kata dilakukan untuk menghindari perhitungan dimensi kata yang melebar. Perhitungan dimensi kata

akan melebar jika kata disingkat atau salah eja tidak diubah karena kata tersebut memiliki kontribusi dalam mempresentasikan dokumen tetapi akan dianggap sebagai entitas yang berbeda proses penyusunan matriks.

5. *Stopword Removal* yaitu proses penghapusan kata-kata yang terdapat pada *stoplist*. *Stoplist* itu sendiri berisi kosakata-kosakata yang bukan merupakan ciri dari suatu dokumen atau kalimat yang tidak penting.

6. *Stemming* ialah suatu proses pemetaan dan penguraian berbagai bentuk (variants) dari suatu kata menjadi bentuk kata dasarnya (*stem*). *Stemming* bertujuan untuk menghilangkan berbagai imbuhan seperti prefiks, sufiks, maupun konfiks yang ada pada setiap kata.

### 2.2.7. Term Frequency Inverse Document Frequency (TF-IDF)

*Term Frequency Inverse Document Frequency* (TF-IDF) ialah suatu metode yang digunakan untuk pembobotan dan menentukan ukuran keterkaitan kata (term) terhadap dokumen dengan memberikan bobot setiap kata [24]. TF-IDF ialah pembobotan dimana bobot suatu kata berdasarkan frekuensi dokumen terbalik yaitu suatu kata yang semakin banyak muncul pada banyak dokumen, maka kata tersebut memiliki bobot yang lebih kecil [34]. Pembobotan TF-IDF dapat dilihat pada persamaan (2.1).

$$W_{ij} = tf_{ij} \times \log \frac{N}{n_j} \dots \dots \dots (2.1)$$

Keterangan:

$W_{ij}$  : Bobot kata  $j$  pada dokumen  $i$

$tf_{ij}$  : Frekuensi kata  $j$  pada dokumen  $i$

$N$  : Jumlah dokumen

$n_j$  : Jumlah dokumen yang terdapat kata  $j$

Apabila persamaan (2.1) terdapat nilai  $N$  sama dengan  $n_j$  maka bobot yang diperoleh adalah 0, maka dapat menggunakan persamaan (2.2)

$$W_{ij} = \log (tf_{ij} + 1) \times \log \frac{N+1}{n_j} \dots \dots \dots (2.2)$$



### 2.2.8. *Naive Bayes Classifier* (NBC)

Algoritma *Naive bayes* ialah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas. *Naive bayes* menggunakan cabang dari matematika yang dikenal dengan teori probabilitas untuk mencari peluang terbesar dari kemungkinan klasifikasi, dengan cara melihat frekuensi tiap klasifikasi pada data training [35].

Metode *Naive bayes classifier* terdapat dua tahap dalam proses klasifikasi teks, yaitu tahap pelatihan dan tahap klasifikasi. Pada tahap pelatihan dilakukan proses yaitu analisis terhadap sampel dokumen berupa pemilihan *vocabulary*, yaitu kata yang mungkin muncul dalam koleksi dokumen sampel hingga dapat serta menjadi representasi dokumen. Selanjutnya adalah penentuan probabilitas bagi tiap kategori berdasarkan sampel dokumen. Pada tahap klasifikasi ditentukan nilai kategori dari suatu dokumen berdasarkan term yang muncul dalam dokumen yang diklasifikasi [36]. Rumus untuk *Naive bayes classifier* pada persamaan berikut [15].

$$P(c_i) = \frac{fd(c_i)}{|D|} \dots \dots \dots (2.3)$$

Keterangan:

- $P(c_i)$  : Menentukan probabilitas  $c_i$  yang merupakan kategori kelas
- $fd(c_i)$  : Jumlah dokumen  $c_i$
- $|D|$  : Jumlah data latih/dokumen

$$P(w_k|c_i) = \frac{f(w_{ki},c_i)+1}{P(c_i)+|W|} \dots \dots \dots (2.4)$$

Keterangan:

- $P(w_k|c_i)$  : Peluang kemunculan kata-kata pada sebuah kategori/kelas,  $w_k$  adalah kata yang muncul pada sebuah kategori.
- $f(w_{ki}, c_i)$  : Nilai kemunculan kata  $w_{ki}$  pada kelas  $c_i$
- $f(c_i)$  : Jumlah keseluruhan kemunculan kata pada kelas  $c_i$
- $|W|$  : Jumlah semua kata dari semua kategori

### 2.2.9. *Support Vector Machine* (SVM)

*Support Vector Machine* ialah suatu metode klasifikasi untuk data linier dan nonlinier. Sebuah SVM merupakan algoritma yang bekerja menggunakan pemetaan nonlinier untuk mengubah data pelatihan asli menjadi dimensi yang lebih tinggi. Dalam dimensi baru ini, ia mencari *hyperplane* yang memisahkan optik linear yaitu, batas keputusan memisahkan tuple dari satu kelas dari kelas yang lain. Dengan pemetaan nonlinier yang tepat ke dimensi yang cukup tinggi,

data dari dua kelas selalu dapat dipisahkan oleh *hyperplane*. SVM berusaha menemukan *hyperplane* menggunakan vektor dukungan esensial atau pelatihan tuples dan margin (ditentukan oleh vektor dukungan) [17].

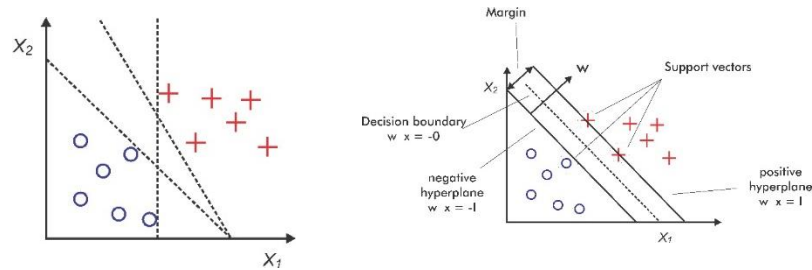
Teknik ini termasuk dalam metode klasifikasi jenis terpandu (*supervised*) karena memiliki target pembelajaran tertentu. Klasifikasi dilakukan dengan mencari *hyperplane* atau garis pembatas (*decision boundary*) yang memisahkan antara satu kelas dengan kelas lainnya. Dalam konsep ini, SVM berusaha untuk mencari *hyperplane* terbaik diantara fungsi yang tidak terbatas jumlahnya. Fungsi yang tidak terbatas dalam pencarian *hyperplane* di metode *Support Vector Machine* merupakan sebuah keuntungan, dimana pemrosesan pasti akan selalu bisa dilakukan bagaimanapun data yang dimilikinya [17]. Berikut ini merupakan kekuatan atau kelebihan dari *Support Vector Machine* (SVM) yaitu [37]:

- 1) Mempunyai kemampuan generalisasi yang tinggi.
- 2) Mampu menghasilkan model klasifikasi yang baik meskipun dilatih dengan himpunan data yang relatif sedikit hanya dengan pengaturan parameter yang sederhana. SVM memiliki konsep dan formulasi yang jelas dengan sedikit parameter yang harus diatur.
- 3) Relatif mudah diimplementasikan karena penentuan SVM dapat dirumuskan dalam masalah QP (*Quadratic Programming*).

Sementara itu, kelemahan atau kekurangan yang terdapat dalam *Support Vector Machine* (SVM) sebagai berikut:

- 1) Sulit diaplikasikan untuk himpunan data dengan jumlah sampel dan dimensi yang sangat besar.
- 2) Umumnya hanya diformulasikan untuk menyelesaikan masalah klasifikasi dua kelas. Walaupun dapat dikembangkan untuk menyelesaikan masalah klasifikasi multi kelas, namun masing-masing strategi multi kelas SVM juga memiliki kelemahan.

Dalam diruang kerja berdimensi tinggi, akan dicari *Hyperplane* yang dapat memaksimalkan margin antara dua kelas. *Hyperplane* klasifikasi Linear SVM dinotasikan sebagai berikut [38]:

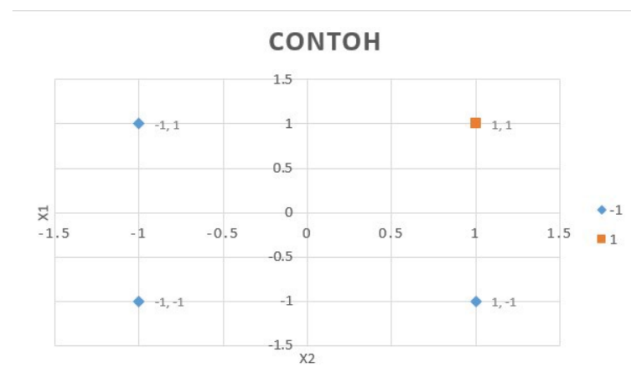


**Gambar 2.1** Hyperplane yang memisahkan dua kelas positif (+1) dan negatif (-1)

Pada gambar 2.1 menjelaskan bahwa *hyperplane* berada di tengah-tengah antara dua kelas, artinya jarak antara *hyperplane* dengan objek-objek data berbeda-beda dengan kelas yang berdekatan yang diberi tanda lingkaran dan positif. Pada metode SVM objek data terluar yang paling dekat dengan *hyperplane* disebut *support vector*. *Support vector* berada di posisi tumpang tindih (overlap) dengan kelas lain yang menyebabkan sifat yang kritis maka hanya *support vector* inilah yang diperhitungkan untuk menemukan *hyperplane* yang paling optimal. Untuk menentukan *hyperplane* langkah awal yang dilakukan membuat tabel yang berisikan data seperti contoh tabel 2.2.

**Tabel 2.2** Contoh Data

Y	X1	X2
1	1	1
-1	1	-1
-1	-1	1
-1	-1	-1



**Gambar 2.2** Plot Contoh Data

Pada gambar 2.2 menjelaskan bahwa terdapat dua kelas yang terdiri -1 yang ditunjukkan dengan warna biru dan 1 ditunjukkan warna *orange*. Pada

masing-masing titik tersebut digunakan untuk mencari pemisah antara data positif dan negatif. Kemudian melakukan penyelesaian menggunakan persamaan linear.

Sedangkan secara matematika, rumus SVM untuk kasus klasifikasi linear adalah [39]:

$$\min \frac{1}{2} \|w\|^2$$

Dimana  $y_i(wx_i + b) \geq 1, i = 1, \dots, l$ . ..... (2.5)

Dan rumus  $w \cdot x + b \geq 1$  ..... (2.6)

Keterangan:  $x_i$  merupakan data masukan,  $i$  merupakan jumlah data,  $y_i$  merupakan keluaran. Berikut adalah persamaan SVM Linear [40]:

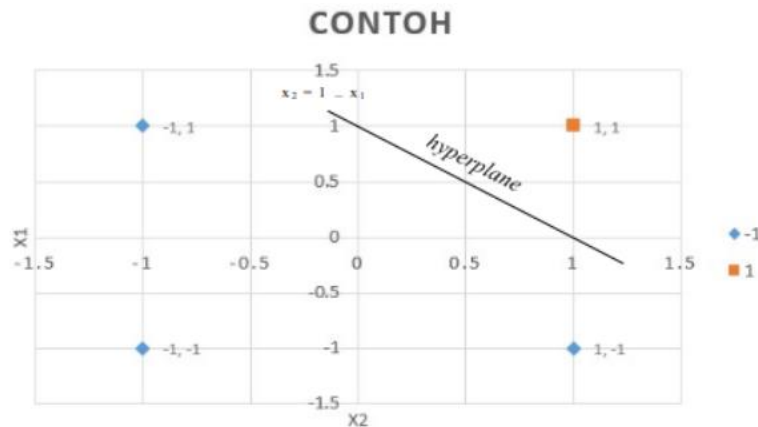
$$w \cdot x - b = 0 \dots\dots\dots (2.7)$$

Keterangan:

$w$ = parameter *hyperplane* yang dicari

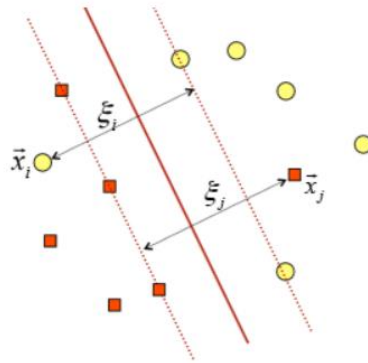
$x$ = data input SVM

$b$ = parameter *hyperplane* yang dicari (bias)



**Gambar 2.4 Menentukan *Hyperplane***

Berdasarkan gambar 2.4 menyatakan bahwa kedua kelas terpisah oleh *hyperplane* secara sempurna. Apabila dua kelas tidak terbelah secara sempurna maka hal itu disebabkan constraint pada persamaan 1 tidak terpenuhi, sehingga optimasi tidak terpenuhi dilakukan [38]. Untuk mengatasi masalah ini, maka dirumuskan kembali dengan teknik *soft margin*.



**Gambar 2.5 Soft Margin**

Pada berbagai kasus, terdapat data yang tidak bisa dilakukan klasifikasi menggunakan metode linear SVM sehingga dikembangkan fungsi *kernel* atau string kernel yang berguna untuk klasifikasi data ke dalam bentuk non linear. Berikut ini adalah beberapa fungsi kernel yang umum digunakan yaitu [41]:

a. Kernel Linear:  $K(x, x_i) = x \cdot x_i \dots \dots \dots (2.8)$

b. *Polinomial*:  $K(x, x_i) = (x_i \cdot x + 1)^p \dots \dots \dots (2.9)$

c. *Radial Basis Function*:  $K(x, x_i) = \exp\left(\frac{-\|x-x_i\|^2}{2\sigma^2}\right) \dots \dots \dots (2.10)$

d. *Sigmoid Kernel*:  $K(x, x_i) = \tanh(kx \cdot x_i - \delta) \dots \dots \dots (2.11)$

Keterangan :  $p$  : pangkat (degree of ),  $\sigma$  : nilai sigma, dan  $\delta$  : nilai delta

Parameter pada metode SVM yaitu seperti parameter  $\gamma$  (*gamma*),  $\sigma$  (*sigma*),  $C$  (*complexity*),  $d$  (*degree*) dan  $\lambda$  (*lambda*) [42]. Parameter yang dibutuhkan tergantung dengan penggunaan kernel yang dipilih dan data yang digunakan agar menghasilkan nilai optimal [43]. Pada penelitian tersebut parameter  $d$  berpengaruh pada kernel *Polinomial*, dan parameter  $\sigma$  berpengaruh pada kernel RBF.

Pada penelitian ini menggunakan *library sklearn Support Vector Classification*. Parameter yang dibutuhkan pada penelitian ini adalah parameter  $C$  (*complexity*),  $d$  (*degree*),  $\sigma$  (*sigma*) dan lainnya.

### 2.2.10. Confusion matrix

*Confusion matrix* juga sering disebut *error matrix*. Pada dasarnya *confusion matrix* memberikan informasi perbandingan hasil klasifikasi yang dilakukan oleh sistem (model) dengan hasil klasifikasi sebenarnya. *Confusion matrix* berbentuk tabel matriks yang menggambarkan kinerja model klasifikasi pada serangkaian data uji yang nilai sebenarnya diketahui [44]. Beberapa *performance matrix* dari *confusion matrix* yang populer adalah *accuracy*, *precision*, *recall* dan *f1 score*.

*Accuracy* menggambarkan seberapa akurat model dapat mengklasifikasikan dengan benar. Maka, *accuracy* merupakan rasio prediksi benar (positif dan negatif) dengan keseluruhan data. Dengan kata lain, *accuracy* merupakan tingkat kedekatan nilai prediksi dengan nilai aktual (sebenarnya). Nilai *accuracy* yang hasil berupa persentase diperoleh dengan persamaan yaitu pada (2.12).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \dots\dots\dots (2.12)$$

*Precision* ialah menggambarkan keakuratan antar model yang diminta dengan hasil prediksi yang diberikan oleh model dan merupakan rasio prediksi benar positif dari seluruh hasil yang di prediksi positif. Nilai *precision* yang hasil berupa persentase dengan persamaan yaitu (2.13).

$$Precision = \frac{TP}{TP+FP} \times 100\% \dots\dots\dots (2.13)$$

*Recall* ialah menggambarkan keberhasilan model dalam menemukan kembali informasi dan merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Nilai *recall* dapat diperoleh persamaan (2.14)

$$Recall = \frac{TP}{TP+FN} \times 100\% \dots\dots\dots (2.14)$$

*F1 Score* ialah perbandingan rata-rata presisi dan recall yang dibobotkan. Nilai *f1 score* dapat diperoleh persamaan (2.15)

$$F1\ Score = 2 \times \frac{Recall * Precision}{Recall + Precision} \dots\dots\dots (2.15)$$

**Tabel 2.3 Confusion Matrix**

		Actual Values	
		1(Positive)	0(Negative)
Predicted Values	1(Positive)	TP	FP
	0(Negative)	FN	TN

Keterangan untuk **Tabel 2.3** Confusion matrix yang menggambarkan nilai:

- *True Positive* (TP), merupakan data positif yang diprediksi memang positif (benar).  
Misal pada kasus *cyber harassment* pada data tersebut mengandung kata kunci *harassment* serta setelah di prediksi memang benar memiliki tujuan untuk melakukan pelecehan atau *harassment*.
- *True Negative* (TN), merupakan data negatif yang diprediksi memang negatif (benar). Misalnya pada data tidak mengandung kata kunci *harassment* lalu setelah di prediksi benar tidak memiliki tujuan untuk melakukan pelecehan atau *harassment*.
- *False Positive* (FP), merupakan data negatif namun diprediksi sebagai data positif (salah atau error). Misalnya pada data tidak mengandung kata kunci *harassment* namun setelah di prediksi mengandung pelecehan atau *harassment*.
- *False Negative* (FN), merupakan data positif namun diprediksi sebagai data negatif (salah atau error). Misalnya pada data mengandung kata kunci *harassment* namun setelah di prediksi tidak memiliki tujuan untuk melakukan pelecehan atau *harassment*.

### 2.2.11. Python

*Python* salah satu bahasa pemrograman tinggi yang bisa melakukan eksekusi dengan sejumlah instruksi multi guna secara langsung (interpretatif) dengan metode *Object Oriented Programming* dan juga menggunakan semantik dinamis untuk memberikan tingkat keterbacaan syntax [45]. Perbedaan *python* dengan bahasa pemrograman lain ialah pada aturan penulisan kode program. *Python* merupakan bahasa pemrograman yang berorientasi obyek dinamis yaitu dapat digunakan oleh bermacam-macam perangkat lunak serta *python* sangat cocok hampir semua sistem operasi seperti *Windows, Linux,*

dan lain-lain [46]. Mempunyai kode yang sederhana dan mudah di implementasikan ialah salah satu kelebihan *Python*.

Berbagai macam *tools* yang dapat digunakan menggunakan bahasa pemrograman *python* seperti *Anaconda*, *Jupyter Notebook* [47], dan *Google Collaboratory* [48]. *Anaconda* ialah distribusi yang bersifat *free* dan *open-source* bahasa pemrograman Python dan R untuk Scientific Computing seperti *Data Science*, *Machine Learning application*, *large scale data processing* *predictive analytics* dan lain sebagainya [47]. Mayoritas *Data Science* menggunakan *Jupyter Notebook* karena sifat nya mudah digunakan, fleksibel (berbentuk web-application) dan tentunya *open-source* [47]. *Google Collaboratory* atau Google Colab merupakan *tools* yang berbasis cloud dan free untuk penelitian. Pada dasarnya Google Colab sama dengan *Jupyter Notebook* atau bisa dikatakan *Jupyter Notebook* yang dijalankan secara online dan gratis [48]. Pada penelitian ini dapat menggunakan semua *tools*, namun lebih mengutamakan dengan menggunakan *Jupyter Notebook* yang sudah dimiliki oleh peneliti.