# BAB II

# KAJIAN LITERATUR

## 2.1. Tinjauan Pustaka

### 2.1.1. *State of The Art*

Pada suatu penelitian, tentu saja sangat penting dalam memahami perkembangan terkini dalam bidang yang diteliti agar dapat mengetahui letak penelitian yang sedang dilakukan dibandingkan dengan penelitian sebelumnya. Tujuan dari kajian literatur ini adalah untuk memberikan gambaran terkait penelitan-penelitian terdahulu yang relevan, metode yang digunakan, serta hasil yang telah dicapat. Penelitian-penelitian terdahulu yang telah dilakukan serta relevan dengan peneliatian yang akan dilakukan dijelaskan pada Tabel 2.1 berikut.

Tabel 2.1 *State of The Art*

| No | Penulis | Data | Metode | Hasil Penelitian |
|---|---|---|---|---|
| 1 | Zilvi Azus Sriyanti, et. al (2024) [15] | Data berasal dari platform twitter (Sekarang X) dengan jumlah data sebanyak 11.483, kemudian dilakukan *preprocessing* data sehingga menjadi 10.256, dengan data dilakukan penyeimbangan antara label positif dan negatif | indoBERT | Hasil penelitian dengan menggunakan BERT ini terbukti dapat memberikan akurasi sebesar 85%, recall 85%, dan F1-score sebesar 85% dengan sentimen positif lebih banyak ditemukan. |
| 2 | Hendra Poltrak, et.al (2024) [17] | Data yang didapatkan bersumber dari media sosial X dengan jumlah data yang diambil sebanyak 72 data | BERT | Penelitian menggunakan model BERT dengan hasil sentimen positif sebesar 8%, negatif sebesar 67%, dan netral sebesar 25%. |
| 3 | Intan Nurma Yulita, et.al (2023) [18] | Data diperoleh dari komentar YouTube yang berkaitan dengan kebijakan pemerintah dalam membatasi perjalanan selama libur Idul Fitri | BERT (IndoBERT) | Hasil penelitian menunjukkan F1-score untuk negatif sebesar 91%, netral sebesar 76% dan positif sebesar 86%, lebih unggul |

| No | Penulis | Data | Metode | Hasil Penelitian |
|---|---|---|---|---|
| | | 2021, dengan total data sebanyak 10.397. Data positif sebanyak 1616, negatif 6.640, dan netral sebanyak 2.141 data. | | dibandingkan dengan *Naïve Bayes* sebesar 74,10%, SVM 81,00% dan LSTM sebesar 82,00% |
| 4 | Dimas Samodra Bimaputra, Edi Sutoyo (2023) [19] | Data yang digunakan adalah data ulasan pelanggan hotel di Bali dari situs *TripAdvisor* berjumlah 3.419 data dengan menggunakan *web scraping*. Data diklasifikasi berdasarkan aspek keselamatan, kebersihan, keamanan, dan layanan dengan sentimen positif dan negatif. | BERT dan Aspect-Based Sentiment Analysis (ABSA) | Hasil dari penelitian menunjukkan bahwa akurasi sentimen umum: 95%, akurasi aspek kebersihan 99%, akurasi aspek kenyamanan 99%, akurasi aspek layanan 95%, dan akurasi aspek keselamatan 100%. |
| 5 | Serly Setyani, et. al (2023) [20] | Data yang digunakan berupa ulasan pengguna aplikasi Bibit pada Google Play Store dalam rentang waktu Januari 2019 hingga Desember 2022 dengan jumlah data sebanyak 3.004 data, dengan data dibagi menjadi 3 kategori, kualitas sistem, kualitas layanan, dan keputusan | IndoBERT dan Aspect-Based Sentiment Analysis (ABSA) | Hasil evaluasi model menunjukkan akurasi aspek layanan sebesar 92% mayoriats ulasan netral, aspek kepuasan pengguna sebesar 87% lebih banyak ulasan positif, dan aspek sistem sebesar 75% dengan di dominasi sentimen netral |
| 6 | Niluh Putu Vania Dyah Saraswati, et. al (2023) [21] | Data bersumber dari repositori GitHUB yang sebelumnya telah melalui *Focus Group Discussion* (FGD) dengan pihak Direktorat Tindak Pidana Siber | BERT | Hasil Penelitian menunjukkan akurasi model dalam klasifikasi perundungan siber mencapai 81% untuk data validasi dan uji, presisi |

| No | Penulis | Data | Metode | Hasil Penelitian |
|---|---|---|---|---|
| | | Bareskrim Polri. Data ini dibagi menjadi 3 kategori, netral, bahasa kasar, dan ujaran kebencian, dengan semua data berbahasa Indonesia. | | dengan 80,67% pada data validasi dan 80,33% pada data uji, serta *recall* 80,33%, dan *F1-score* mencapai 80,67% untuk validasi dan 80,33% untuk data uji |
| 7 | Daffa Fadilah Putra, et.al (2024) [22] | Data berasal dari platform media sosial X dengan kata kunci Pemilihan Presiden 2024, dengan total data sebanyak 5000 *tweet.* Data dibagi tiga label, positif 2129 data, negatif 1480 data, dan netral 1388 data, data tambahan sebanyak 18.000 *twee*t digunakan untuk prediksi dukungan terhadap tiga kandidat. | BERT (*bert-base-uncased*) dan CNN | Hasil penelitian menunjukkan akurasi model BERT mencapai 90,02% lebih baik dibanding CNN sebesar 88,19%. Prediksi dukungan terhadap kandidat presiden berdasarkan sentimen anaisis, Prabowo Subianto mendapat 43,82% dukungan, Ganjar Pranowo mendapatkan 33,83% dukungan, Anies Baswedan mendapatkan 22,35% dukungan. |
| 8 | Ali Areshey, Hassan Mathkour (2023) [23] | Data bersal dari *Yelp Review Dataset* berisi ulasan pengguna tentang layanan dan produk dengan 10.00 ulasan, 5000 positif dan 500 negatif. | BERT (*bert-base-uncased*) | Model BERT mencapai akurasi 97,3% lebih baik dibandingkan SVM sebesar 94,6%, serta KNN dan NB sebesar 85,5% dan 91,1%. |
| 9 | Marlon Santiago Viñan-Ludeña, et. al (2022) [24] | Data diperoleh dari 90.725 unggahan Instagram dan 235.755 *tweet* Twitter yang membahas destinasi wisata di Granada, Spanyol, yang terdiri dari | BERT (Spanish-BERT) | Model BERT untuk bahasa Spanyol akurasinya sebesar 75,7%, Tweeteval untuk bahasa inggris sebesar 75,6%. Destinasi wisata paling sering |

| No | Penulis | Data | Metode | Hasil Penelitian |
|---|---|---|---|---|
| | | bahasa inggris dan spanyol. | | disebut oleh wisatawan meliputi Alhambra, Albaicín, dan Sierra Nevada, dengan mayoritas sentimen positif. Sentimen negatif terhadap harga tiket masuk yang mahal, sistem pemesanan tiket yang buruk, dan kualitas layanan yang rendah. |
| 10 | Mrityunjay Singh, et.al (2021) [25] | *Dataset* terdiri dari 592.838 *tweet* global dan 3.946 *tweet* dari India, dikumpulkan dari 20 Januari 2020 hingga 25 April 2020 menggunakan Twitter API dan *Twitter Scraper.* | BERT-*base* | Hasil analisis sentimen pada global 26,1% netral, 17,4% negatif ringan, 14,5% negatif kuat, 8,3% positif ringan, dan 9,0% positif kuat. Hasil analisis sentimen di India 22,9% netral, 15,9% negatif ringan, 14,9% negatif kuat, 9,4% positif ringan, dan 10,6% positif kuat. |
| 11 | Jiawei, wu et.al [26] | Data dari unggahan media sosial Twitter dan Facebook, dengan dianotasi menjadi positif, negatif, dan netral. | BERT-*base* | Hasil studi BERT lebih unggul dibanding dengan metode tradisional, dengan akurasi 89% meningkat 12% dibanding metode terbaik dari *machine learning*, berhasil mengatasi bahasa informal, sarkasme, dan konteks ambigu dibanding model SVM, LSTM, CNN. |

Berdasarkan dari pemaparan penelitian terdahulu pada Tabel 2.1 tersebut, sehingga disimpulkan bahwa penggunaan analisis sentimen sangat diperlukan

dalam memperoleh opini dari teks. Informasi yang berasal dari berbagai platform perlu dilakukan analisis sentimen sehingga dapat diketahui makna dan emosi dari opini yang sedang dibahas.

Penelitian-penelitian terdahulu tersebut membahas berbagai komentar maupun ulasan terkait isu-isu yang sedang terjadi maupun terhadap suatu produk, tempat, ataupun jasa. Adanya penelitan analisis sentimen ini dapat memberikan peluang untuk meningkatkan pengetahuan bagaimana opini publik beredar. Dalam pengolahan data teks, analisis sentimen diharuskan juga dapat mengenali subjek, polaritas, serta pemegang opini, sehingga analisis sentimen dapat menjadi solusi dalam memahami opini publik.

Penelitian terdahulu seperti yang dilakukan oleh Zilvi Azus Sriyanti, et. al [15] serta Poltak et.al [17] telah menggunakan BERT dalam analisis sentimen media sosial X, akan tetapi cakupan dari analisis masih terbatas pada klasifikasi sentimen secara umum tanpa mempertimbangkan aspek demografi. Studi lain seperti Yulita et.al [18] juga telah menggunakan IndoBERT dalam menganalisis komentar terkait kebijakan pemerintah, tetapi tidak mengaitkannya dengan hasil sentimen dengan faktor demografis. Begitu juga studi yang dilakukan oleh Daffa Fadilah Putra et.al [22] berfokus pada sentimen politik, serta studi yang dilakukan Saraswati et.al [21] meneliti tentang ujaran kebencian. Dari penelitan-penelitian tersebut, sehingga menunjukkan bahwa BERT memiliki akurasi yang tinggi dalam mengklasifikasi sentimen.



Penelitian terdahulu tersebut memberikan pemaparan bahwa BERT dapat memiliki efektivitas yang baik dalam melakukan analisis sentimen, dengan ditunjukkan memiliki akurasi yang tinggi serta dapat mengenali klasifikasi sentimen dengan baik, akan tetapi tidak mengaitkan hasil klasifikasi sentimen tersebut dengan faktor demografis. Maka dari itu penelitian ini memanfaatkan algoritma BERT dalam melakukan analisis sentimen terhadap opini publik tentang isu *"married is scary"* pada media sosial *X*. Dengan memvisualisasikan sentimen berdasarkan faktor demografis, sehingga dapat memberikan wawasan lebih mendalam terkait isu pernikahan di masyarakat berdasarkan demografi.

## 2.2. Landasan Teori

### 2.2.1. Analisis Sentimen

Analisis sentimen atau sering disebut sebagai *Opinion Mining* merupakan suatu studi komputasi mengenai opini, sikap, maupun emosi orang lain terhadap suatu hal, baik itu perorangan, peristiwa maupun topik tertentu, dengan topik-topik tersebut berwujud dalam bentuk ulasan [27]. Sentimen analisis juga dapat dikatakan sebagai proses untuk mengekstraksi emosi ataupun pendapat dari suatu teks yang berisikan topik tertentu [28]. Tujuan dari sentimen analisis ini adalah untuk mengetahui opini ataupun sikap penulis terhadap suatu topik, sehingga mendapatkan pengetahuan dari pesan yang terdapat dari ulasan tersebut [27].

Agar mengetahui pesan yang terkandung dari suatu teks yang dilakukan oleh publik, maka dapat dilakukan dengan menggunakan sentimen analisis. Dalam mengetahui pesan tersebut, maka harus mengetahui terlebih dahulu bahwa teks tersebut apakah subjektif atau objektif, karena hanya teks subjektif ternyata yang mengandung sentimen, sedangkan teks objektif hanya mengandung informasi faktual saja. Kemudian, teks subjektif tersebut diklasifikasikan menjadi tiga kategori berdasarkan dari sentimen yang terdapat dalam teks, seperti positif, negatif, maupun netral [28].

Pada sentimen analisis ini terdapat tingkatan sentimen yang berguna untuk menunjukkan berbagai tingkatan yang terkandung dari sentimen, seperti tingkatan dokumen (*document level*) yang di dalamnya hanya diberikan polaritas tunggal saja seperti positif, negatif, ataupun objektif. Selain itu, tingkatan kalimat (*sentence level*), dimana di dalam dokumentasi diklasifikasikan ke dalam tingkat kalimat, dengan setiap kalimat di analisis terpisah serta diklasifikasin dengan negatif, positif, maupun objektif. Terakhir tingkatan frasa *(phase level*), dimana merupakan analisis yang lebih mendalam dengan berhubungan terhadap frasa maupun aspek lainnya dalam sebuah kalimat, yang kemudian diklasifikasikan menjadi positif, negatif maupun objektif [28].

### 2.2.2. Pernikahan

Dalam kehidupan, pernikahan merupakan salah satu hal kebutuhan mendasar pada kehidupan manusia yang mencakup aspek biologis, sosial, dan emosional. Sebagai institusi yang telah ada sejak awal peradaban, pernikahan tidak hanya

berfungsi sebagai pemenuhan naluri dasar manusia, akan tetapi juga berperan sebagai pembentukan struktur sosial yang teratur[29]. Melalui pernikahan terbentuklah keluarga, yang merupakan unit dari paling dasar dalam masyarakat serta berkintribusi dalam proses regenerasi, pendidikan nilai moral, serta perlindungan terhadap individu. Selain memberikan kelangsungan keturunan, pernikahan juga menghadirkan rasa aman, kasih sayang, serta kestabilan emosi yang penting bagi kesejahteraan hidup seseorang [30].

Di Indonesia, pernikahan memiliki dasar hukum formal melalui Undang-Undang Nomor 1 Tahun 1974 tentang Perkawinan, yang kemudian diperbarui melalui Undang-Undang Nomor 16 Tahun 2019. Regulasi tersebut mencerminkan bahwa negara memandang pernikahan sebagai pilar penting dalam membentuk keluarga dan menjaga ketertiban sosial, sekaligus menjamin perlindungan hukum dan keadilan bagi warga negara yang memasuki kehidupan pernikahan [31].

Namun, di tengah modernisasi dan perubahan nilai-nilai generasi muda di Indonesia, makna institusi pernikahan mulai mengalami pergeseran. Pada generasi muda terutama Gen Z, menunjukkan sikap yang semakin kritis terhadap konsep pernikahan tradisional. Dimana pada saat ini muncul rasa ketakutan terhadap pernikahan, dengan memandang pernikahan terhadap ketidakpastian ekonomi, beban sosial, dan pengalaman negatif keluarga menjadi alasan utama mengapa sebagian anak muda memilih menunda atau bahkan menolak pernikahan. Selain itu, ekspektasi gender yang tidak seimbang juga memengaruhi persepsi negatif terhadap pernikahan, terutama di kalangan perempuan [30].

Fenomena ini juga diperkuat oleh hasil survei *Harvard Youth Poll* pada tahun 2025 yang dilaporkan oleh *New York Post*, di mana hanya 57% responden muda di Amerika Serikat yang menyatakan ingin menikah suatu hari nanti. Survei tersebut juga menunjukkan bahwa anak muda lebih memprioritaskan hubungan romantis jangka panjang tanpa ikatan hukum, serta kehidupan pribadi yang stabil secara emosional dan finansial. Temuan ini menandakan adanya perubahan orientasi nilai, di mana pernikahan tidak lagi dianggap sebagai tujuan utama kehidupan, melainkan sebagai pilihan yang dipertimbangkan secara rasional dan pragmatis sesuai kondisi individu [32].

### 2.2.3. Aplikasi X (Twitter)

Aplikasi X atau sebelumnya dikenal dengan Twitter merupakan suatu platform yang memberikan penggunanya untuk dapat berbagi pesan (*tweets*) hingga 2080 karakter. Platform ini memiliki peran yang sangat besar dalam komunikasi global, khususnya dalam penyebaran informasi secara langsung, baik itu pembahasan terkait diskusi publik, pemasaran digital, maupun isu sosial yang sedang marak terjadi [33].

Platform ini selain dapat menjadi sumber informasi serta diskusi terkait beberapa topik tersebut, dimana pengguna dapat mengikuti akun-akun yang identik dengan minat serta dapat berinteraksi dengan konteks yang terdapat dalam aplikasi ini. Platform ini juga selain dapat digunakan sebagai sumber komunikasi, tetapi dalam konteks penelitian juga dapat digunakan dalam melakukan analisis sentimen, pemetaan tren, maupun sebagai alat untuk pemerintah maupun organisasi dalam memahami opini publik yang beredar. Sehingga dengan analisis sentimen dapat mengetahui opini serta pandangan publik terkait isu yang beredar di masyarakat [33].

### 2.2.4. *Natural Language Processing* (NLP)

Pemrosesan bahasa alami atau sering disebut dengan *Natural Language Processing (NLP)* merupakan suatu cabang dari kecerdasan buatan (*Artificial Intelligent*) yang berfokus kepada interaksi antara komputer dengan bahasa manusia. Penggunaan NLP ini mencakup dengan berbagai teknik untuk memahami, memproses, serta menghasilkan suatu bahasa alami secara otomatis, termasuk seperti pemrosesan teks, analisis sentimen, penerjemahan mesin, maupun pengenalan suara. Guellil et.al menyakan bahwa NLP sendiri memiliki peran yang dapat membantu manusia dalam berbagai bidang seperti pencarian informasi, *chatbot*, maupun analisis *big data*. Hal ini karena NLP menggabungkan teknik linguistik komputasional dengan model *deep learning* dalam meningkatkan pemahaman serta intepretasi bahasa manusia yang dilakukan oleh komputer [34].

*Natural Language Processing (NLP)* ini tersusun dalam beberapa komponen, seperti *dictionary/leksion* yang berguna sebagai rujukan yang menjelaskan makna dari kata-kata, dimana setiap *entry* data dalam kamus mewakilkan informasi tentang semua kata. Selain itu ada *parser* yang berupa elemen yang paling menentukan

dalam NLP, dimana kerjanya dengan mengidentifikasi tiap kata yang kemudian dibuat peta kata-kata tersebut dalam struktur disebut pohon *parser* yang berisi makna dalam semua kata dan cara untuk menggabungkan semua kata tersebut. Kemudian sistem representasi pengetahuan, yakni dengan menganalisis keluaran untuk menentukan maknanya. Berikutnya, *output translator* yang merupakan terjemahan yang menjelaskan sistem pengetahuan serta melakukan kegiatan yang dapat berupa jawaban dari bahasa alami maupun hal khusus sesuai dengan program komputer lainnya [35].

### 2.2.5. *Text Preprocessing*

*Text preprocessing* merupakan suatu tahapan awal dalam pemrosesan *Natural Language Processing* (NLP) yang bertujuan untuk mengubah teks yang tidak terstruktur menjadi suatu format yang lebih dapat dipahami oleh mesin. Penggunaan teknik ini dibutuhkan dalan konteks NLP seperti analisis sentimen [36]. Data teks secara umum memiliki struktur kalimat yang tidak baik serta memiliki banyak *noise* sehingga menjadikan informasi di dalam teks tersebut tidak dapat diekstraksi secara langsung, sehingga dibutuhkanlah suatu *text preprocessing* untuk menanganinya.

Dengan melakukan *text preprocessing* sehingga data yang akan digunakan menjadi lebih bersih, relevan, serta efisien untuk digunakan dalam pemrosesan lebih lanjut. *Text preprocessing* ini berguna dalam mengubah bentuk data yang asalnya belum terstruktur menjadi data yang terstruktur sesuai dengan kebutuhan dalam proses *mining* seperti analisis sentimen [35].

### 2.2.6. *Tokenization*

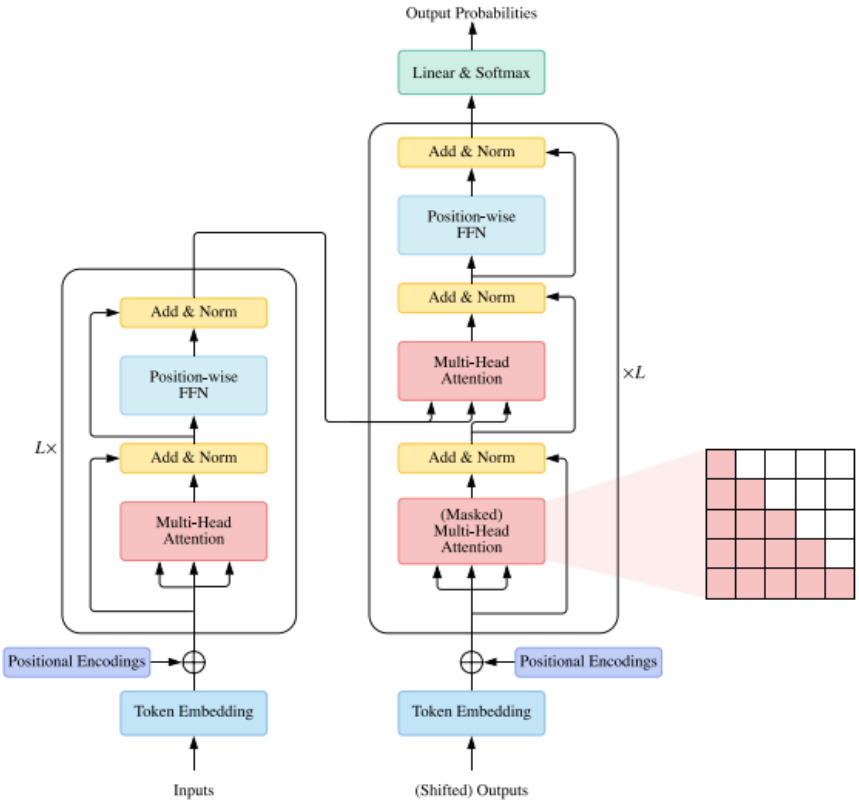
*Tokenization* merupakan suatu proses untuk membagi teks menjadi unit-unit yang lebih kecil, seperti kata, frasa, maupun kalimat. Dalam langkah ini merupakan langkah penting karena dapat membantu dalam pemetaan teks menjadi elemen yang dapat diproses lebih lanjut [35]. Contoh data yang telah dilakukan *tokenization* seperti pada Tabel 2.2 berikut.

Tabel 2.2 Tokenisasi Teks

| **Teks** | **Hasil *Tokenization*** |
|---|---|
| “apakah culo dan boyo bermain bola di depan rumah boyo?” | “culo”, “dan”, “boyo”, “bermain”, “bola”, “di”, “depan”, “rumah” |

### 2.2.7. *Transformers*

*Transformers* merupakan suatu model dari *deep learning* yang dikenal dalam berbagai bidang seperti NLP, pemrosesan suara, maupun *computer vision*, dengan awal penggunaannya hanya untuk *sequence-to-sequence* pada *machine translation* saja. *Transformers* ini mamanfaatkan mekanisme *self-attention* dalam memahami hubungan antara kata dalam suatu teks [37]. Arsitektur dari *transformers* dapat dilihat pada Gambar 2.1 berikut.




Gambar 2.1 Arsitektur *Transformers*[38]

Dalam memahami terhadap suatu teks, *transformers* dapat menghitung representasi kata yang akan muncul berikutnya dalam kata tertentu, kemudian akan dibandingkan dengan setiap kata lain pada kalimat. Skor dari perbandingan tersebut dapat menentukan besar atau tidaknya kontribusi pada setiap kata lainnya terhadap representasi berikutnya. Kemudian skor tersebut akan digunakan sebagai bobot untuk setiap kata yang tertimbang dari semua representasi kata yang terdapat dalam jaringan yang terhubung penuh dengan kata yang dibandingkan tersebut [39].

*Transformers* memanfaatkan jaringan saraf untuk menerjemahkan setiap kata dalam teks dengan berisi *encoder* sebagai bagian yang berfungsi untuk membaca

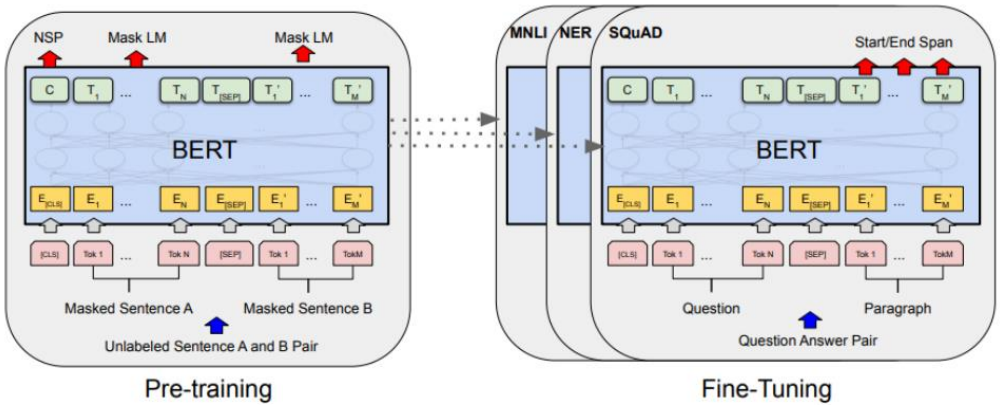
kalimat yang masuk serta menghasilkan representasinya. Kemudian *decoder* akan menghasilkan kalimat keluaran kata per kata dengan merujuk kepada representasi yang dihasilkan oleh *encoder* [39].

Proses *transformers* diawali dengan menghasilkan representasi awal atau *embedding* pada setiap kata, selanjutnya dengan menggunakna *self-attention*, informasi-informasi dari semua kata akan digabungkan sehingga dapat menghasilkan representasi baru perkata yang terdapat dalam teks [39].

### 2.2.8. *Bidirectional Encoder Representations from Transformers* (BERT)

BERT merupakan suatu model bahasa yang rilis pada tahun 2018 oleh Google dimana berbasis *transformers* [15]. BERT berguna dalam meningkatkan pemahaman dari teks dalam *natural language processing*. Proses pemahaman BERT dilakukan dengan menggunakan *bidirectional*, yakni memahami konteks dari teks dari sebelumnya dan setelahnya pada suatu kalimat [40].

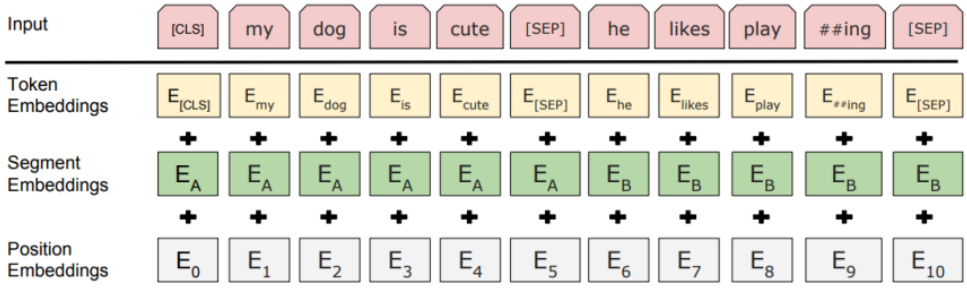
Penggunaan BERT pada saat ini telah banyak digunakan diberbagai aspek seperti analisis sentimen, klasifikasi teks, maupun pencarian informasi [40]. Arsitektur dari BERT ini merupakan arsitektur jaringan yang telah dilatih di *dataset* yang besar dengan berbagai banyak bahasa. Sehingga pada proses pelatihan pada setiap lapisan BERT yang telah terdapat bobot-bobot yang baik. Pelatihan tersebut membutuhkan tambahan *layer* yang digunakan untuk proses klasifikasi, dimana *layer*-nya tersebut terdiri atas teks, *pre-processing* BERT *encoder*, *droupout*, serta *classifier*. Pertama kali dirilis BERT memiliki dua ukuran, yakni BERT-*base* dengan 12 *layer*, 768 ukuran *vector*, dan 110 juta parameter. Kedua adalah BERT-*large* yang memiliki 24 *layer*, 1024 ukuran *vector*, dan 340 juta parameter [41]. Penjelasan mengenai tahapan kerja *pre-training* dan *fine-tuning* pada BERT dapat dilihat pada Gambar 2.2 berikut.




Gambar 2.2 *Pre-Training* dan *Fine-Tuning* BERT[15]

Pada Gambar 2.2 merupakan tahapan kerja pada BERT, yang terdiri atas dua tahapan kerja, yakni *pre-training* dan *fine-tuning*. Pada *pre-training* model dilatih dengan data yang tidak berlabel, dengan mengikutsertakan model untuk memahami bahasa serta konteks melalui *Masked Laguage Modeling* (MLM), yakni sebagian kata dalam teks disembunyikan dan model harus menebaknya, serta dengan *Next Sentence Prediction* (NSP), dimana model memprediksi apakah dua kalimat memiliki hubungan atau tidak. MLM berguna untuk menggabungkan konteks dari dua arah kiri dan kanan, sedangkan NSP berguna untuk menggabungkan representasi dari pasangan teks pada *pre-training*. Sedangkan *fine-tuning* yakni menggunakan BERT yang telah di *pre-train* untuk digunakan dalam proses menyelesaikan masalah [15]. Representasi input dalam BERT dijelaskan pada Gambar 2.3 berikut.






Gambar 2.3 Representasi *Input* dari BERT [15]

Gambar 2.3 merupakan reperesentasi input dalam BERT, dimana dimulai dengan *token embedding*, dengan setiap urutan dilalui dengan *classification token* (CLS) serta token khusus *separator token* (SEP) yang berguna untuk memisahkan setiap kalimat. Setiap kalimat dibedakan antara satu dan lainnya dikarenakan

memiliki *segment embedding*, sedangkan posisi setiap token dalam urutan dilakukan oleh *positional embedding*. Sehingga input dari BERT merupakan penggabungan dari *token embedding, segment embedding* serta *positional embedding* [15].

### 2.2.9. *Confusion Matrix*

*Confusion matrix* merupakan suatu instrumen yang berguna untuk melakukan evaluasi sistem klasifikasi dengan menghasilkan suatu tabel matriks perbandingan hasil dari prediksi model terhadap kelas sebenarnya pada data. Kegunaan dari matriks ini adalah untuk menghitung dari hasil prediksi seperti *accuracy, precision, recall,* serta *F1 score* [15].

Pada *confusion matrix* terdiri atas *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) serta *False Negative* (FN). Dimana dari hal tersebut berguna untuk mengukur *accuracy, precision, recall*, dan *F1-score* [15].

*Precision* bertujuan untuk mengukur banyaknya prediksi positif yang sebenarnya diprediksi dengan benar [42]. *Precision* dinyatakan dengan rumus persamaan 2.1 berikut:

$$Precision = \frac{TP}{TP + FP} \tag{2.1}$$

Sedangkan *accuracy* berguna untuk mengukur ketepatan model dalam memprediksi kelas-kelas yang benar, dimana prediksi positif dan negatif dibandingkan dengan total jumlah data [42]. Berikut rumus dari *accuracy* pada persamaan 2.2 berikut.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.2}$$

*Recall* berguna untuk mengukur tingkat banyaknya dari semua kelas positif yang telah diidentifikasi benar oleh model [42]. Berikut rumus *recall* dinyatakan dalam persamaan 2.3.

$$Recall = \frac{TP}{TP + FN} \tag{2.3}$$

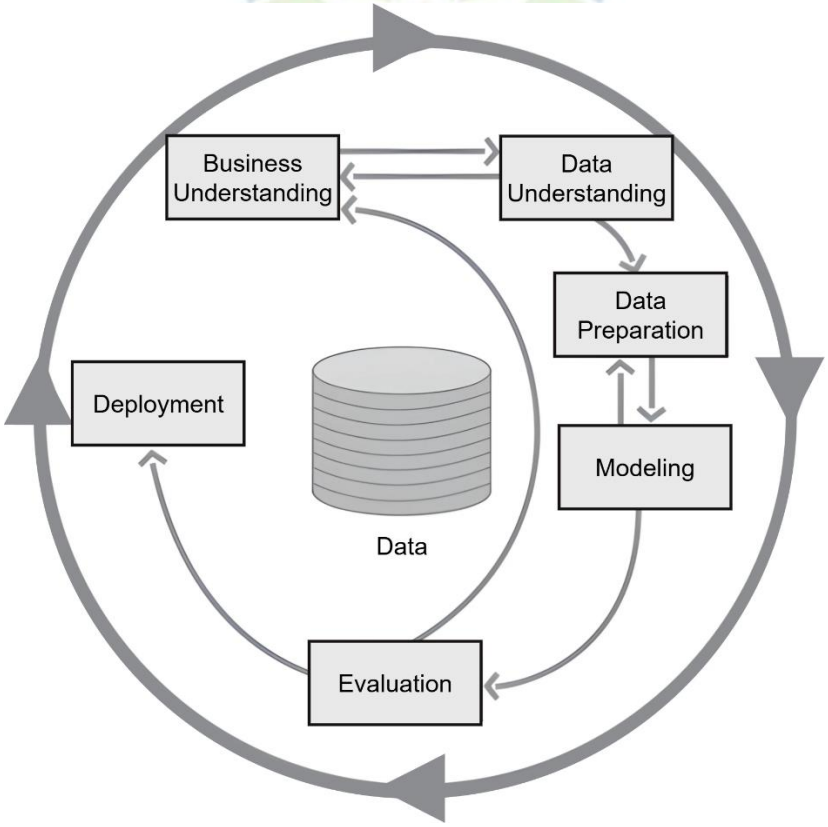
*F1-score* berguna untuk membandingkan rata-rata dari *precision* dengan *recall*. Hal ini berguna untuk mengetahui agar model apakah seimbang atau tidak dalam

memprediksi antara *recall* dan *precission*, karena tidak boleh ada yang mendominasi salah satunya [42]. Berikut rumus dari *F1-score* pada persamaan 2.4.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (2.4)$$

### 2.2.10. *Cross Industry Standard Process for Data Mining* (CRISP-DM)

CRISP-DM merupakan suatu metodologi yang dikembangkan pada tahun 1990-an untuk proyek-proyek *data science* serta *data mining*. Model ini dirancangan untuk memberikan referensi yang memberikan struktur dalam pelaksanaan proyek, kemudian meningkatkan pemahaman serta komunikasi antara pemangku kepentingan, dan menjadikan proyek lebih terorganisis serta dapat diulang dengan mudah [43]. Penjelasan mengenai alur dari CRISP-DM dijelaskan pada Gambar 2.4 berikut.




Gambar 2.4 Fase CRISP-DM[44]

Pada gambar 2.4 tersebut merupakan fase-fase yang terdapat dalam CRISP-DM. Dalam CRISP-DM terdapat enam fase yang dilakukan. Enam fase tersebut terdiri atas [43].

1. ***Business Understanding***

Pada fase ini bertujuan untuk memahami tujuan serta kebutuhan proyek dari prespektif bisnis. Pada tahap ini, permasalah yang diinginkan diselesaikan melalui analisis data diidentifikasi dengan jelas agar solusi yang dihasilkan relevan dengan tujuan bisnis.

2. ***Data Understanding***

Fase ini berfokus pada pengumpulan data awal dalam memperoleh pemahaman awal terkait data tersebut. Setelah data dikumpulkan, kemudian diperiksa serta dianalisis untuk dipahami karakteristik dari data serta mengetahui pola dalam data. Selain itu, bertujuan untuk mengetahui dari kualitas data apakah terdapat informasi yang hilang atau tidak konsisten sehingga dapat memengaruhi hasil analisis. Dengan pemahaman yang baik pada data, model dapat lebih akurat serta dapat memenuhi kebutuhan.

3. ***Data Preparation***

Pada tahap ini, data yang akan digunakan dalam pemodelan dipersiapkan dengan melakukan seleksi, transformasi, dan pembersihan data. Persiapan data ini dilakukan mencakup seperti pembersihan data yang tidak lengkap, normalisasi agar skala data seragam, serta *encoding* untuk mengubah data kategori menjadi format numerik yang dipahami oleh mesin.

4. ***Modeling***

Fase ini melibatkan pemilihan teknik pemodelan yang sesuai, setelah data sudah siap. Selain itu pada tahap ini dilakukan pengaturan parameter untuk menghasilkan model yang optimal. Jika ditemukan masalah dalam *dataset*, maka dapat kembali ke fase sebelumnya untuk penyesuaian.

5. ***Evaluation***

Model yang telah dikembangkan dievaluasi untuk memastikan bahwa model tersebut sesuai dengan tujuan bisnis yang telah ditetapkan serta memiliki kinerja model yang baik. Tahap ini berguna untuk meninjau langkah-langkah yang telah dilakukan untuk memastikan keakuratan model. Evaluasi model ini dapat dilakukan dengan menggunakan metrik tertentu, seperti akurasi, *presision, recall,* serta *F1-score*. Model yang telah dilatih dicek performanya, apabila tidak memenuhi standar maka perlu dilakukan perbaikan. Tahapan ini memastikan agar model tidak hanya

menghasilkan prediksi yang baik pada saat pelatihan, juga memiliki prediksi yang baik dalam skenario di dunia nyata.

**6. *Deployment***

Tahap terakhir dalam CRISP-DM adalah implementasi model ke dalam sistem bisnis operasional. Model yang telah dievaluasi dan disetujui kemudian diterapkan dalam proses bisnis untuk memberikan nilai tambah bagi organisasi. Implementasi bisa dalam bentuk integrasi dengan sistem manajemen data yang sudah ada atau pengembangan dashboard analitik untuk mendukung pengambilan keputusan. Setelah diterapkan, model harus dipantau secara berkala untuk memastikan bahwa kinerjanya tetap optimal dan dapat diperbarui sesuai dengan perubahan dalam data atau lingkungan bisnis.