

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Selama beberapa tahun terakhir, telah terjadi perubahan populasi yang cukup signifikan di berbagai belahan dunia, ditandai dengan penurunan tingkat fertilitas global dan munculnya fenomena isu *childfree*, yaitu sebuah keputusan perjanjian yang dibuat secara individu ataupun pasangan untuk tidak memiliki anak [1]. Laporan *World Population Prospects 2022* dari PBB mencatat penurunan angka kelahiran global dari 3,2 anak perwanita pada tahun 1990 menjadi 2,3 anak perwanita pada tahun 2021, dan diproyeksikan turun hingga 2,1 anak perwanita pada tahun 2050 [2].

Isu *childfree* tidak hanya menjadi pilihan pribadi, tetapi juga mencerminkan perubahan nilai-nilai dalam masyarakat modern. Survei *Pew Research Center* pada tahun 2021 mengungkapkan bahwa 44% masyarakat berusia 18-49 tahun yang belum memiliki anak menyatakan mereka “mungkin tidak” atau “pasti tidak” akan memiliki anak di masa depan, meningkat 7% dari tahun 2018 [3]. Fenomena *childfree* ini dipengaruhi oleh berbagai faktor, seperti aspirasi karir, kebebasan pribadi, kekhawatiran terhadap ketidakstabilan finansial, dan kesadaran terhadap keberlanjutan lingkungan. Studi dari *Michigan State University* mengungkapkan bahwa alasan utama yang mendorong keputusan ini mencakup 60% berkeinginan untuk memiliki lebih banyak waktu luang, 50% ingin lebih fokus pada pencapaian karir, dan 54% terdapat kekhawatiran tentang ketidakstabilan finansial [4].

Dalam konteks demografi, peningkatan jumlah individu atau pasangan yang memilih hidup tanpa anak dapat berkontribusi pada percepatan penuaan populasi (*aging population*) dan berkurangnya jumlah penduduk usia produktif [5]. Kondisi ini dapat memengaruhi dinamika ekonomi suatu negara, termasuk ketersediaan tenaga kerja, beban sistem pensiun, hingga ketahanan ekonomi dalam jangka Panjang [6]. Selain itu, pilihan hidup *childfree* dapat memengaruhi aspek-aspek lainnya, seperti nilai-nilai budaya, persepsi tentang peran gender, dan stabilitas sosial [7].

Di negara-negara barat seperti Amerika Serikat, diskusi tentang *childfree* cenderung lebih terbuka dan diterima sebagai pilihan gaya hidup yang valid. Sebuah studi oleh *National Center for Health Statistics AS* menunjukkan bahwa tingkat kesuburan di Amerika Serikat telah menurun menjadi 1,64 anak per wanita pada tahun 2020, jauh di bawah angka 2,1 yang diperlukan untuk mempertahankan jumlah populasi yang stabil, karena angka yang dianggap normal atau stabil untuk menjaga populasi suatu negara tetap seimbang umumnya adalah sekitar 2,1 anak per wanita [8].

Sementara itu, di negara-negara dengan nilai-nilai tradisional yang kuat seperti Indonesia, konsep *childfree* masih sering dipandang kontroversial dan bertentangan dengan norma sosial. Meskipun data spesifik tentang *childfree* di Indonesia terbatas, Badan Pusat Statistik (BPS) mencatat penurunan *Total Fertility Rate* (TFR) dari 2,41 anak per wanita pada tahun 2010 menjadi 2,18 pada tahun 2020 [9].

Bersamaan dengan perubahan pola kehidupan masyarakat terkait isu *childfree*, perkembangan pesat teknologi komunikasi internet telah menimbulkan pengaruh signifikan terhadap masyarakat untuk berkomunikasi serta mengakses informasi secara mudah dan cepat di berbagai portal termasuk media sosial, yang menciptakan ruang diskusi, seperti memberikan pujian, kritik, hingga bentuk ekspresi lainnya [10]. Salah satu *platform* media sosial yang sering dimanfaatkan untuk diskusi publik adalah X (sebelumnya Twitter) [11]. X telah berkembang menjadi ruang digital di mana pengguna dari seluruh dunia dapat langsung berdiskusi, bertukar pendapat, dan mengutarakan pendapat termasuk tentang isu *childfree*.

Penggunaan media sosial, termasuk X, cukup signifikan di berbagai negara, termasuk di Indonesia dan Amerika Serikat. Menurut hasil riset *We Are Social* pada bulan Januari tahun 2024, jumlah pengguna media sosial di Indonesia mencapai 191,4 juta jiwa dari jumlah populasi sebesar 277,7 juta jiwa (68,9% dari populasi) [12]. Sedangkan di Amerika Serikat, jumlah pengguna media sosial mencapai 246 juta jiwa dari total jumlah populasi sebesar 334 juta jiwa (74% dari populasi) [13]. Data ini menunjukkan *platform* media sosial memiliki peran yang signifikan dalam

kehidupan masyarakat antar kedua negara, menjadikannya sumber data yang kaya untuk memahami opini publik, termasuk tentang isu *childfree*.

Namun demikian, volume dan kompleksitas data yang dihasilkan dari interaksi di media sosial menciptakan tantangan dalam menganalisis opini publik secara akurat dan komprehensif. Di sinilah peran analisis sentimen menjadi krusial. Analisis sentimen, sebuah cabang dari pemrosesan bahasa alami, memungkinkan peneliti untuk mengekstrak dan mengkuantifikasi sikap dan emosi yang diekspresikan dalam teks digital secara otomatis [14].

Dalam beberapa tahun terakhir, kemajuan dalam bidang kecerdasan buatan telah membuka peluang baru untuk analisis sentimen yang lebih canggih. Salah satu terobosan paling signifikan adalah pengembangan algoritma *Bidirectional Encoder Representations from Transformers* (BERT). BERT telah menunjukkan peningkatan akurasi yang signifikan dalam tugas-tugas NLP, termasuk analisis sentimen, terutama dalam menangkap nuansa dan konteks dalam bahasa yang digunakan di media sosial [15]. Terdapat berbagai jenis model BERT yang populer, namun tugas akhir ini menerapkan model BERT *base multilingual*.

Pemanfaatan BERT *base multilingual*, memungkinkan analisis sentimen dapat dilakukan secara lebih efektif karena mampu memahami konteks serta menangani ambiguitas dalam teks multibahasa. Model ini memiliki keunggulan dalam menangkap makna dan konteks kalimat, memahami pengecualian, serta mengenali ekspresi dalam teks. BERT *base multilingual* telah melalui proses *pre-training* pada 104 bahasa, termasuk bahasa Indonesia dan bahasa Inggris [16]. Dibandingkan dengan model BERT *monolingual* atau model lainnya, BERT *base multilingual* memiliki keunggulan khusus dalam menganalisis data dwibahasa karena dilatih secara bersamaan pada kedua bahasa target, sehingga dapat menangkap pola lintas bahasa dengan lebih akurat. Sesuai dengan penelitian yang dilakukan oleh Kuncahyo Setyo Nugroho dengan nilai akurasi keberhasilan 80% [17].

Pemilihan Indonesia dan Amerika Serikat dalam studi ini didasarkan pada perbedaan karakteristik budaya dan sosial yang cukup kontras terkait isu *childfree*. Indonesia, sebagai negara dengan nilai-nilai tradisional, budaya kekeluargaan yang kuat, serta norma sosial yang masih sangat menjunjung tinggi pernikahan dan keturunan, cenderung memandang keputusan *childfree* sebagai sesuatu yang

kontroversial dan kurang diterima secara luas. Sebaliknya, Amerika Serikat merepresentasikan masyarakat Barat modern yang lebih individualistik, dengan tingkat penerimaan sosial yang lebih tinggi terhadap pilihan gaya hidup alternatif, termasuk keputusan untuk tidak memiliki anak. Perbedaan ini menawarkan peluang analisis yang menarik untuk melihat bagaimana latar belakang budaya memengaruhi persepsi masyarakat terhadap isu *childfree*, sehingga memungkinkan perbandingan sentimen yang lebih komprehensif dan bermakna antara kedua negara.

Berdasarkan latar belakang tersebut, tugas akhir ini bertujuan untuk menganalisis sentimen komparatif antara masyarakat Indonesia dan Amerika Serikat terhadap isu *childfree*, menggunakan data dari media sosial X serta memanfaatkan kemampuan canggih algoritma *Bidirectional Encoder Representations from Transformers* (BERT) dengan model *BERT base multilingual*. Analisis akan berfokus pada identifikasi dan perbandingan proporsi sentimen positif, negatif, dan netral di kedua negara yang memiliki latar belakang budaya berbeda. Oleh karena itu, tugas akhir ini berjudul “Analisis Sentimen Isu *Childfree* di Indonesia dan Amerika Serikat Menggunakan *BERT (Bidirectional Encoder Representations from Transformers)*”.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, dapat dirumuskan masalah utama pada tugas akhir ini, yaitu:

1. Bagaimana menerapkan algoritma BERT dengan model *BERT base multilingual* untuk analisis sentimen di Indonesia dan Amerika Serikat terhadap isu *childfree* di media sosial X?
2. Bagaimana kinerja yang dihasilkan dari penerapan model *BERT base multilingual* dalam melakukan analisis sentimen yang mengandung konteks dwibahasa?
3. Bagaimana hasil analisis sentimen terhadap isu *childfree* di Indonesia dan Amerika Serikat?

## 1.3 Tujuan

Adapun tujuan dilakukannya tugas akhir ini adalah sebagai berikut:

Mengetahui cara menerapkan algoritma BERT dalam model *bert base multilingual* untuk analisis sentimen di Indonesia dan Amerika Serikat terhadap isu *childfree* di media sosial X.

Mengetahui kinerja model BERT *base multilingual* dalam melakukan analisis sentimen yang mengandung konteks dwibahasa antara lain Bahasa Indonesia dan Bahasa Inggris.

Membandingkan hasil proporsi sentimen positif, negatif dan netral terhadap isu *childfree* antara Indonesia dan Amerika Serikat dan menganalisis perbandingan sentimen.

#### **1.4 Manfaat**

Penulisan tugas akhir ini diharapkan mampu memberikan manfaat sebagai berikut:

##### **1.4.1 Bagi Penulis**

Penelitian ini akan memberikan manfaat bagi penulis, di antaranya:

- a. Memberikan pengalaman dan pengetahuan baru bagi penulis dalam melakukan penelitian ilmiah, khususnya dalam bidang analisis sentimen, pemrosesan bahasa alami (NLP), dan penggunaan model BERT.
- b. Menghasilkan karya ilmiah yang dapat dijadikan sebagai referensi untuk penelitian selanjutnya terkait analisis sentimen dan isu sosial di media sosial.

##### **1.4.2 Bagi Masyarakat**

Penelitian ini akan memberikan manfaat bagi masyarakat atau pengguna, di antaranya:

- a. Memberikan wawasan mengenai persepsi publik terkait isu *childfree* di Indonesia dan Amerika Serikat berdasarkan analisis media sosial X.
- b. Membantu masyarakat dalam memahami tren opini publik serta faktor-faktor yang mempengaruhi sentimen terhadap isu *childfree* di kedua negara.
- c. Memberikan informasi berbasis data yang dapat digunakan oleh masyarakat, organisasi, dan pihak terkait dalam merancang strategi komunikasi dan kebijakan yang lebih baik terkait isu keluarga dan sosial.

##### **1.4.3 Bagi Akademik**

Penelitian ini akan memberikan manfaat bagi akademik atau di lingkungan pendidikan, di antaranya:

- a. Memberikan kontribusi dalam pengembangan ilmu pengetahuan dan teknologi, khususnya dalam bidang analisis sentimen berbasis *deep learning* dan pemrosesan bahasa alami (NLP).
- b. Menambah referensi ilmiah yang relevan dan terkini tentang penggunaan model BERT dalam analisis sentimen isu sosial di media sosial.
- c. Menjadi acuan dan inspirasi bagi peneliti lain yang akan melakukan penelitian sejenis atau lanjutan mengenai analisis sentimen, perbandingan opini publik lintas negara, serta pengaruh media sosial dalam membentuk persepsi masyarakat terhadap isu tertentu.

### **1.5 Batasan Masalah**

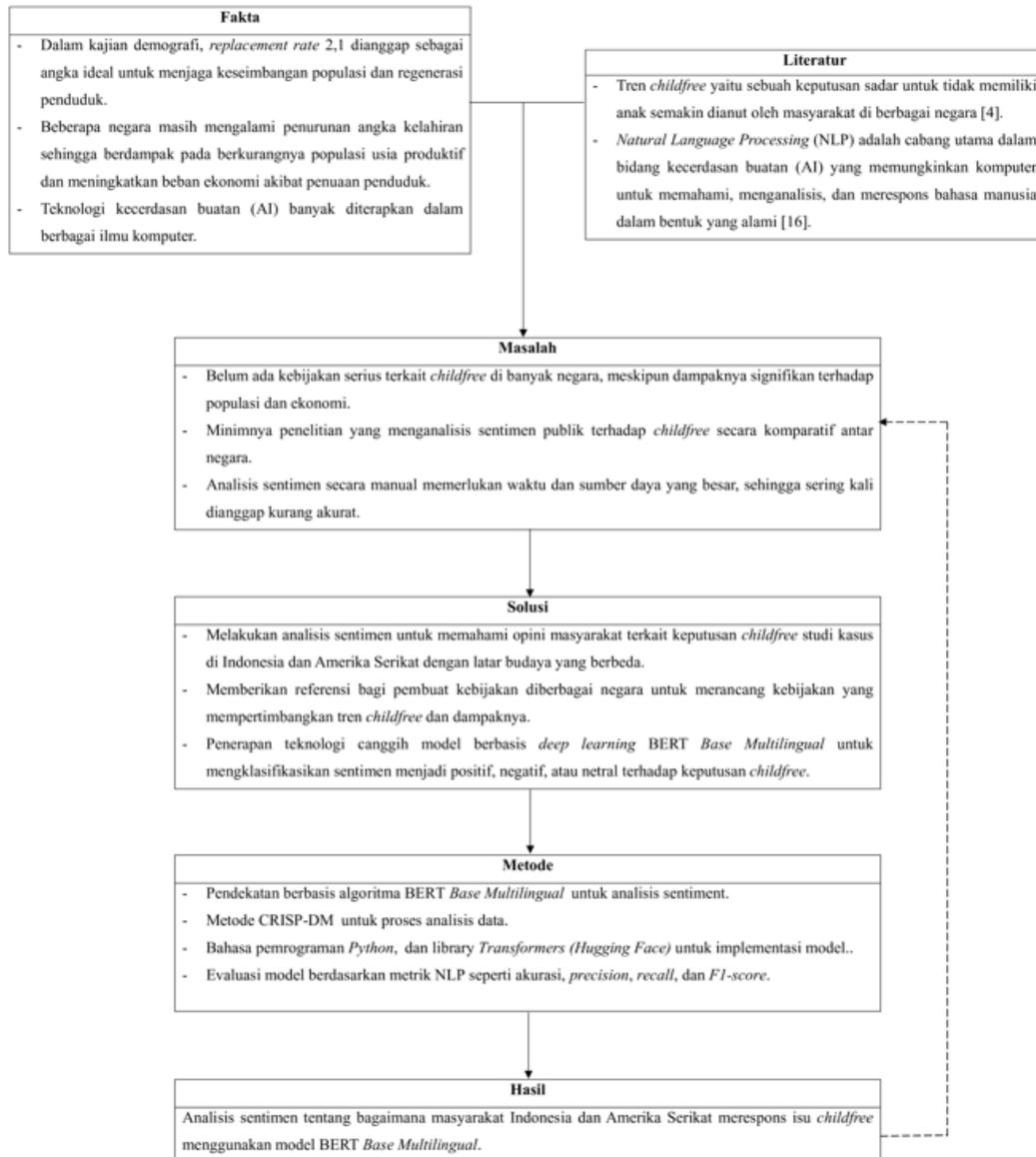
Tugas akhir ini akan dibatasi pada beberapa hal:

1. Data yang digunakan berasal dari media sosial X, yang dipublikasikan pada rentang periode tanggal 1 Januari 2024 hingga 31 Desember 2024, dan dikumpulkan menggunakan *tweet harvest* yang memanfaatkan autentikasi sesi pengguna (*Auth Token*).
2. Tugas akhir ini membatasi data pada maksimum 13.000 dataset dari kedua negara yang relevan berdasarkan kata kunci tertentu, yaitu "*childfree*".
3. Tweet yang dianalisis terdiri dari tweet berbahasa Indonesia untuk data masyarakat Indonesia dan tweet berbahasa Inggris untuk data masyarakat Amerika Serikat, yang telah difilter dan diidentifikasi berdasarkan metadata lokasi.
4. Pelabelan data dilakukan secara otomatis, untuk dataset berbahasa Indonesia, menggunakan model *IndoBERTweet-base-Indonesian-sentiment-analysis* dan untuk dataset berbahasa Inggris, menggunakan model *BERTweet-base-sentiment-analysis*.
5. Sentimen dikelompokkan menjadi tiga kategori yakni positif, negatif dan netral.
6. Tugas akhir ini hanya mencakup lima fase dari metodologi CRISP-DM, yaitu: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, dan *Evaluation*. Fase *Deployment* tidak termasuk dalam ruang lingkup tugas akhir ini.

## 1.6 Kerangka Pemikiran

Dalam tugas akhir ini, pertama-tama, informasi faktual dan literatur terkait menjadi landasan yang kritis untuk mengidentifikasi permasalahan yang akan diatasi. Dengan memahami gap dalam pengetahuan yang sudah ada, permasalahan penelitian dapat diuraikan dengan jelas, memberikan konteks mengenai urgensi dan relevansinya. Dalam menghadapi permasalahan tersebut, solusi yang diusulkan kemudian diuraikan, menjelaskan metode atau pendekatan yang akan diambil untuk mengatasi kesenjangan pengetahuan tersebut. Metode kemudian membentuk langkah-langkah sistematis yang akan diambil dalam menjalankan penelitian, termasuk teknik pengumpulan analisis data hingga metode.

Dari implementasi metode tersebut, hasil akhir dari penelitian ini akan dihasilkan, dalam hal ini akan menghasilkan model analisis sentimen berbasis BERT *Base Multilingual* untuk membandingkan sentimen di Indonesia dan Amerika Serikat terhadap isu *childfree*, dilengkapi dengan visualisasi hasil analisis. Seluruh proses ini membentuk kerangka penelitian yang saling terkait dan mendukung relevansi dampak yang ditimbulkan. Pada Gambar 1.1 menunjukkan kerangka pemikiran dalam tugas akhir ini.



Gambar 1. 1 Kerangka Pemikiran

Gambar 1.1 menggambarkan hubungan antara setiap tahap dalam tugas akhir ini, dimulai dari identifikasi masalah hingga analisis hasil yang diperoleh. Dengan adanya kerangka pemikiran ini, proses penelitian dapat berjalan secara sistematis dan terarah, memastikan bahwa setiap langkah yang dilakukan menghasilkan temuan yang dapat memberikan wawasan lebih lanjut mengenai persepsi masyarakat terhadap isu *childfree* di dua negara.

### 1.7 Sistematika Penulisan

Pada bagian sistematika penulisan laporan mencakup susunan penulisan laporan tugas akhir dengan menjelaskan isi setiap bab, urutan penulisan, serta

hubungan antar bagian dalam laporan. Berikut sistematika penulisan laporan tugas akhir dalam tugas akhir ini.

## **BAB I PENDAHULUAN**

Dalam bab ini mencakup isi bagian latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, kerangka pemikiran penelitian, serta sistematika penulisan.

## **BAB II KAJIAN LITERATUR**

Bab ini membahas mengenai tinjauan literatur dan penelitian terdahulu, konsep konsep, teori-teori, model, dan rumus yang menjadi landasan dalam proses analisis permasalahan dengan topik masalah yang diambil.

## **BAB III METODOLOGI PENELITIAN**

Bagian bab metodologi penelitian ini berisi penjelasan urutan serta metode yang diterapkan dalam penelitian, diuraikan secara sistematis dan terstruktur.

## **BAB IV HASIL DAN PEMBAHASAN**

Pada bab ini menjelaskan dua hal utama, pertama adalah pemaparan tentang temuan atau hasil penelitian yang diperoleh berdasakan tahapan penelitian yang telah dilaksanakan. Selanjutnya adalah pembahasan hasil atau temuan penelitian sebagai jawaban terhadap rumusan masalah penelitian.

## **BAB V SIMPULAN DAN SARAN**

Dan pada bab ini berfokus pada penarikan kesimpulan yang diperoleh dari hasil penelitian yang telah dilakukan, menjawab rumusan masalah, serta menyajikan saran untuk penelitian mendatang yang dapat dilakukan agar meningkatkan kualitas dari penelitian tersebut.