

## ***Ensemble Deep STS Siamese Network* untuk Pendeteksian Kontradiksi Peraturan di Indonesia**

**Abstrak** – Indonesia adalah negara hukum yang memiliki peraturan-peraturan dengan tingkatan hirarki tertentu. Peraturan-peraturan baru dibuat karena munculnya ancaman-ancaman baru baik internal maupun eksternal terhadap negara Indonesia. Proses pembuatan peraturan-peraturan baru tersebut tidak mudah, pemerintah memerlukan waktu yang lama untuk memeriksa keselarasan peraturan-peraturan baru tersebut dengan peraturan-peraturan yang sudah ada. Proses pemeriksaan tersebut saat ini masih dilakukan secara manual. Penelitian ini merancang model pemrosesan bahasa alami berbasis *deep learning* untuk melakukan identifikasi keselarasan peraturan tersebut secara otomatis. Kami membuat arsitektur model *Ensemble Deep Semantic Textual Similarity (STS) Siamese Network*. Model ini dibangun dengan menggabungkan hasil-hasil prediksi model tunggal berarsitektur *Siamese Network*. Sebelum melakukan pelatihan model, teks diproses dengan tahapan *lowercasing*, penghapusan penomoran, pengolahan tanda baca, pembakuan ejaan, penghapusan kata umum, pengolahan kata imbuhan, dan tokenisasi. Arsitektur *Siamese Network* terdiri atas empat lapisan, yaitu lapisan *Cross Encoder*, BERT Word Embedding, Pooling, dan Softmax Classifier. Variasi BERT word embedding yang digunakan adalah Sentence IndoBERT dan SimCSE IndoBERT yang dikombinasikan dengan variasi lapisan *mean pooling* dan CLS pooling. Hasil evaluasi menunjukkan efektivitas model dalam mengidentifikasi aturan selaras, kontradiksi, dan netral. Variasi model tunggal terbaik didapatkan oleh model dengan kombinasi arsitektur Sentence IndoBERT dengan CLS pooling, dengan akurasi mencapai 91,12% dan F1-score 90,01%. Penelitian ini menunjukkan peningkatan performa model ketika dilakukan *ensemble* penggabungan empat variasi model, yaitu SIndoBERT-Mean, SIndoBERT-CLS, SimCSEIndoBERT-Mean, dan SimCSEIndoBERT-CLS, dengan nilai akurasi 91,75% dan F1-score 90,68%. Performa model *ensemble* ini memberikan potensi besar dalam efektivitas penggunaan pemrosesan bahasa alami dalam mengidentifikasi aturan yang selaras dan kontradiksi.

**Kata Kunci:** NLP, STS, Siamese Network, *Ensemble*, Hukum

### **I. PENDAHULUAN**

#### **1.1 Latar Belakang**

Indonesia adalah negara hukum yang menjalankan segala urusan pemerintahan berdasarkan hukum. Dalam menjalankan peraturan perundang-undangan, Indonesia menganut sistem hukum *civil law* atau sistem hukum kontinental (Nurhardianto, 2015).

Indonesia memiliki beberapa jenis peraturan. Setiap peraturan tersebut memiliki hirearki. Hirearki ini memastikan bahwa seluruh peraturan yang memiliki hirearki lebih rendah tidak boleh kontradiksi dengan seluruh peraturan yang memiliki hirearki yang lebih tinggi. Susunan hierarki peraturan perundang-undangan dari tertinggi ke terendah berturut-turut, yaitu:

1. Undang-Undang Dasar Negara Republik Indonesia Tahun 1945;
2. Ketetapan Majelis Permusyawaratan Rakyat;
3. Undang-Undang/Peraturan Pemerintah Pengganti Undang-Undang;
4. Peraturan Pemerintah;
5. Peraturan Presiden;
6. Peraturan Daerah Provinsi; dan
7. Peraturan Daerah Kabupaten/Kota.

Pemerintah telah berkomitmen untuk menerapkan tata kelola yang baik dalam menjalankan pemerintahan. Hal ini dijelaskan dalam Rencana Besar Reformasi Birokrasi melalui Peraturan Presiden Republik Indonesia No. 81 tahun 2010, dengan tujuan menciptakan birokrasi pemerintah yang kompeten dan berintegritas tinggi pada tahun 2025. Dalam rangka pemenuhan komitmen ini, penjagaan tata kelola yang baik penting untuk dilakukan dengan memastikan bahwa proses pembuatan peraturan berjalan konsisten sesuai dengan peraturan yang berlaku. Selain itu, harmonisasi antara berbagai tingkatan perundang-undangan juga harus diperhatikan sehingga tercipta keselarasan pada pasal dan ayat dalam peraturan perundang-undangan.

Peraturan perundang-undangan senantiasa mengalami perubahan seiring zaman berkembang karena munculnya ancaman, tantangan, hambatan, dan gangguan (ATHG) baru. Peraturan yang saat ini berlaku perlu untuk disesuaikan dengan berbagai faktor seperti aturan yang gemuk (*hyper regulation*), saling berkontradiksi, saling tumpang tindih, multitafsir dan multiinterpretasi, inkonsistensi, menciptakan beban tak perlu, dan biaya yang tinggi (Sumodiningrat et al., 2022). Menghadapi tantangan ini, diperlukan upaya segera untuk menjaga keselarasan dan keharmonisan dalam sistem peraturan perundang-undangan di Indonesia.

Saat ini terdapat teknik-teknik pemrosesan bahasa alami untuk mencari keselarasan teks. Keselarasan teks tersebut dapat dilakukan dengan membandingkan satu kalimat dengan kalimat lain atau membandingkan satu dokumen utuh dengan dokumen lain. Penelitian yang dilakukan oleh Sinaga & Hansun, (2018) menerapkan algoritma Rabin-Karp untuk klasifikasi kategori keseriusan plagiarisme dokumen berbahasa Indonesia dengan akurasi klasifikasi

yang rendah sebesar 70%. K-Means clustering dan cosine distance juga diterapkan untuk mencari keselarasan antara dua dokumen dan berhasil meraih akurasi 93,33% (Usino et al., 2019). Namun, akurasi tinggi tersebut disebabkan jumlah dokumen yang digunakan hanya ada 15 sehingga model performa model tidak variative ke banyak dokumen. Rajagukguk dan Khodra mengembangkan model VRep + SVM (support vector machine) untuk melakukan klasifikasi similaritas kalimat berbahsa Indonesia dan mendapatkan nilai *F1-score* sebesar 70,37% (Rajagukguk & Leylia Khodra, 2018). Dengan permasalahan yang sama, Nurjaman dkk. melakukan klasifikasi menggunakan CNN (*Convolutional Neural Network*) dengan nilai *F1-score* rendah sebesar 66% (Nurjaman et al., 2020).

Oleh karena itu, kami membuat model *Ensemble Deep STS Siamese Network*, yaitu model *ensemble* penggabungan empat model tunggal dengan asitektur dasar Siamese, yaitu SIndoBERT-Mean, SIndoBERT-CLS, SimCSEIndoBERT-Mean, dan SimCSEIndoBERT-CLS. Arsitektur Siamese digunakan dalam penelitian ini karena arsitektur ini merupakan dasar arsitektur Sentence BERT *Embedding* yang memiliki performa baik dalam permasalahan *semantic textual similarity* (Reimers & Gurevych, 2019). Model ini akan dilatih untuk melihat kontradiksi dan keselarasan peraturan-peraturan dengan membandingkan satu ayat dengan satu ayat lainnya.

## 1.2 Rumusan Masalah

- a. Bagaimana mengidentifikasi aturan baru yang selaras, kontradiksi atau netral dengan aturan yang telah ada dengan memanfaatkan teknologi kecerdasan buatan berbasis *deep learning* dalam bidang pemrosesan bahasa alami?
- b. Bagaimana mengimplementasikan arsitektur model *deep learning* untuk proses identifikasi aturan yang selaras, kontradiksi, dan netral?
- c. Bagaimana melakukan analisis terhadap model *deep learning* terbaik yang telah diimplementasikan untuk proses identifikasi aturan yang selaras, kontradiksi, dan netral?

## 1.3 Tujuan

Penelitian ini bertujuan untuk menelaah beberapa teknik kecerdasan buatan berbasis *deep learning* untuk menyelesaikan permasalahan identifikasi aturan baru yang selaras, kontradiksi atau netral. Penelitian ini juga bertujuan mengimplementasi dan menganalisis beberapa model *deep learning* untuk permasalahan tersebut sehingga didapatkan model *deep learning* terbaik untuk menyelesaikan masalah tersebut.

## 1.4 Manfaat

Bagi pemerintah:

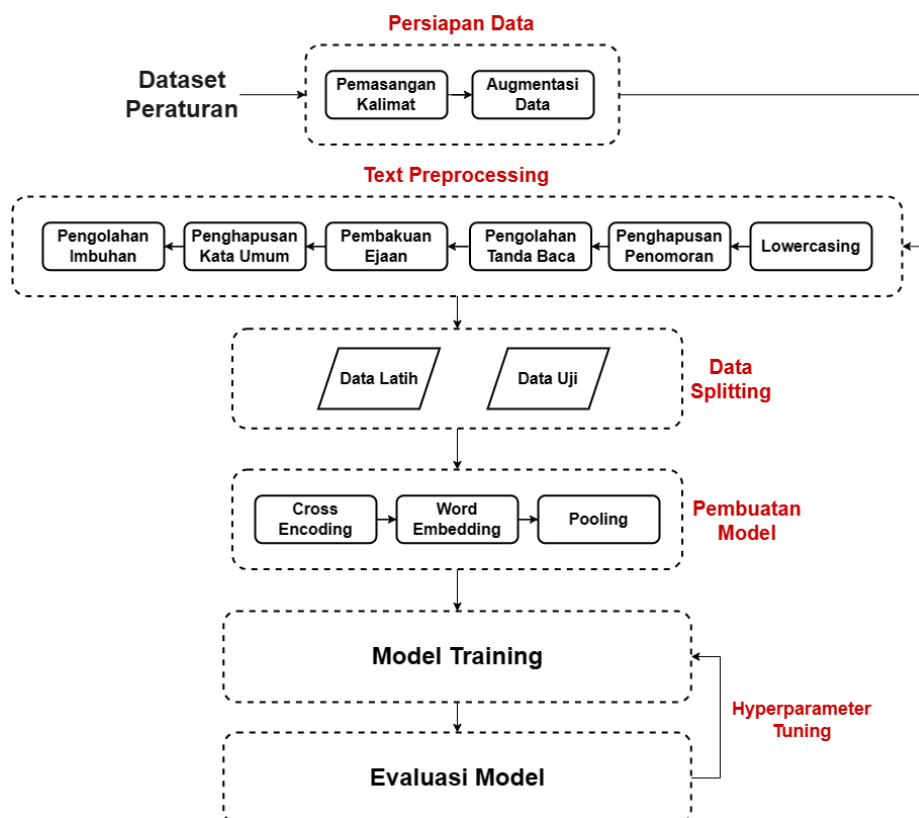
- Pemeriksaan peraturan perundang-undangan untuk mengidentifikasi aturan yang selaras atau kontradiksi yang sebelumnya harus memeriksa ayat dari pasal-pasal secara manual menjadi otomatis.
- Pemerintah lebih cepat dalam mengambil kebijakan seperti, mengubah, mengganti, menghapus, dan revisi terhadap pasal yang teridentifikasi kontradiksi.

Bagi peneliti:

- Peneliti dapat mengembangkan model yang kami ajukan atau mengembangkan model baru dalam permasalahan pendeteksian kontradiksi peraturan Indonesia.
- Peneliti dapat menggunakan dataset yang telah dibuat oleh kami sebagai tolak ukur permasalahan pendeteksian kontradiksi peraturan Indonesia.

## II. METODOLOGI

Penelitian ini dilakukan dalam beberapa tahap, yaitu tahap persiapan data, *text preprocessing*, *data splitting*, *cross encoding*, BERT word embedding, *pooling*, pembuatan model, dan evaluasi model. Diagram alir tahapan pembangunan model identifikasi keselarasan aturan dapat dilihat pada Gambar 1.



**Gambar 1. Diagram Alir Tahapan Pembangunan Model Identifikasi Keselarasan Aturan**

## 2.1 Persiapan Data

Data terdiri atas 3999 pasangan kalimat dengan label kontradiksi, selaras, dan netral. Data dengan label kontradiksi adalah pasangan kalimat yang memiliki makna berlawanan, data dengan label selaras adalah pasangan kalimat yang memiliki makna serupa, dan data dengan label netral adalah pasangan kalimat yang memiliki konteks berbeda. Data terdiri atas 515 pasangan kalimat dari Undang-Undang, 183 pasangan kalimat dari Perpu, 1.535 pasangan kalimat dari Peraturan Pemerintah, 266 pasangan kalimat dari Peraturan Presiden, dan 1.500 pasangan kalimat dari Peraturan Menteri. Rincian banyaknya pasangan tiap dokumen dapat dilihat di Tabel 1.

**Tabel 1. Rincian Banyaknya Pasangan Kalimat Tiap Jenis Dokumen**

Jenis Dokumen	Label	Banyak Data
Undang Undang	Kontradiksi	171
	Selaras	175
	Netral	169
Perpu	Kontradiksi	61
	Selaras	61
	Netral	61
Peraturan Pemerintah	Kontradiksi	209
	Selaras	673
	Netral	653
Peraturan Presiden	Kontradiksi	84
	Selaras	91
	Netral	91
Peraturan Menteri	Kontradiksi	479
	Selaras	542
	Netral	479
<b>Total</b>		3999

### 2.1.1 Pemasangan Kalimat

Pendeteksian kontradiksi peraturan dilakukan dengan membandingkan satu pasang kalimat atau ayat yang ada di dataset. Satu pasangan ayat dapat terdiri atas dua ayat dengan hierarki yang berbeda ataupun hierarki yang sama. Contoh proses pemasangan jenis peraturan dapat dilihat pada Tabel 2.

**Tabel 2. Contoh Proses Pemasangan Jenis Peraturan**

Jenis Pemasangan	Jenis Peraturan Kalimat Pertama	Jenis Peraturan Kalimat Kedua
------------------	---------------------------------	-------------------------------

Hierarki berbeda	Undang Undang	Peraturan Presiden
Hierarki sama	Undang-Undang	Undang-Undang

### 2.1.2 Augmentasi Data

Data dengan label kontradiksi memiliki jumlah yang sedikit sehingga dibutuhkan beberapa jenis teknik augmentasi untuk kalimat kontradiksi. Teknik augmentasi kalimat kontradiksi yang digunakan di antaranya: perubahan kata kerja aktif menjadi kata kerja pasif atau sebaliknya, perubahan objek kalimat, pertukaran subjek dan objek kalimat, perubahan kata keterangan bilangan, penambahan kata tidak atau bukan, dan antonim bentuk terikat. Contoh implementasi masing-masing jenis augmentasi data kontradiksi dapat dilihat pada Tabel 3.

**Tabel 3. Contoh Implementasi Augmentasi Data Kontradiksi**

<b>Jenis Augmentasi</b>	<b>Sebelum Augmentasi</b>	<b>Setelah Augmentasi</b>
Perubahan kata kerja aktif menjadi kata kerja pasif atau sebaliknya	..... Menteri dapat <b>memberhentikan</b> sementara waktu anggota Direksi yang bersangkutan. Jika dalam waktu 2 (dua) bulan setelah .....	..... Menteri dapat <b>diberhentikan</b> sementara waktu oleh anggota Direksi yang bersangkutan. Jika dalam waktu 2 (dua) bulan setelah .....
Perubahan objek	..... <b>Menteri</b> dapat memberhentikan sementara waktu anggota Direksi yang bersangkutan. Jika dalam waktu 2 (dua) bulan setelah .....	..... <b>Presiden</b> dapat memberhentikan sementara waktu anggota Direksi yang bersangkutan. Jika dalam waktu 2 (dua) bulan setelah .....
Pertukaran subjek dan objek kalimat	..... <b>Menteri</b> dapat memberhentikan sementara waktu <b>anggota Direksi</b> yang bersangkutan. Jika dalam waktu 2 (dua) bulan setelah .....	..... <b>anggota Direksi</b> dapat memberhentikan sementara waktu <b>Menteri</b> yang bersangkutan. Jika dalam waktu 2 (dua) bulan setelah .....
Perubahan kata keterangan bilangan	..... Menteri dapat memberhentikan sementara waktu anggota Direksi yang bersangkutan. Jika dalam waktu <b>2 (dua)</b> bulan setelah .....	..... Menteri dapat memberhentikan sementara waktu anggota Direksi yang bersangkutan. Jika dalam waktu <b>4 (empat)</b> bulan setelah .....
Penambahan kata tidak atau bukan	..... Menteri <b>dapat</b> memberhentikan sementara waktu anggota Direksi yang	..... Menteri <b>tidak dapat</b> memberhentikan sementara waktu anggota Direksi yang bersangkutan.

Jenis Augmentasi	Sebelum Augmentasi	Setelah Augmentasi
	bersangkutan. Jika dalam waktu 2 (dua) bulan setelah .....	Jika dalam waktu 2 (dua) bulan setelah .....
Antonim bentuk terikat	Pemerintah dapat memberikan hibah kepada pemerintah daerah dalam rangka rehabilitasi dan rekonstruksi <b>pascabencana.</b>	Pemerintah dapat memberikan hibah kepada pemerintah daerah dalam rangka rehabilitasi dan rekonstruksi <b>prabencana.</b>

Selain data berlabel kontradiksi, data dengan label selaras memiliki jumlah yang sedikit sehingga perlu dilakukan teknik augmentasi kalimat selaras. Teknik augmentasi kalimat selaras yang digunakan antara lain, perubahan kalimat aktif menjadi kalimat pasif atau sebaliknya, perubahan posisi kata keterangan, dan sinonim frasa. Contoh penerapan teknik augmentasi kalimat selaras dapat dilihat pada Tabel 4.

**Tabel 4. Contoh Proses Augmentasi Data**

Jenis Augmentasi	Sebelum Augmentasi	Setelah Augmentasi
Perubahan kalimat aktif menjadi kalimat pasif atau sebaliknya	<b>Pemerintah pusat dapat memberikan hibah kepada pemerintah daerah</b> dalam rangka rehabilitasi dan rekonstruksi pascabencana.	<b>Pemerintah daerah dapat diberikan hibah oleh pemerintah pusat</b> dalam rangka rehabilitasi dan rekonstruksi pascabencana.
Perubahan posisi kata keterangan	Pemerintah dapat memberikan hibah kepada pemerintah daerah <b>dalam rangka rehabilitasi dan rekonstruksi pascabencana.</b>	<b>Dalam rangka rehabilitasi dan rekonstruksi pascabencana,</b> Pemerintah dapat memberikan hibah kepada pemerintah daerah
Sinonim frasa	Pemerintah dapat memberikan hibah kepada pemerintah daerah dalam rangka rehabilitasi dan rekonstruksi <b>pascabencana.</b>	Pemerintah dapat memberikan hibah kepada pemerintah daerah dalam rangka rehabilitasi dan rekonstruksi <b>setelah terjadi bencana.</b>

## 2.2 Text Preprocessing

### 2.2.1 Lowercasing

Seluruh karakter huruf kapital diubah menjadi huruf kecil. Hal ini dilakukan agar model mengenali representasi angka token yang sama antara kata-kata serupa meskipun terdapat perbedaan variasi huruf besar dan kecil. Misalnya, kata “Arsip” dan “arsip” akan

dikenali model sebagai satu token yang sama. Penerapan dari *lowercasing* dapat terlihat pada Tabel 5.

**Tabel 5. Contoh Penerapan *Lowercasing***

Dokumen	Sebelum <i>Lowercasing</i>	Setelah <i>Lowercasing</i>
Peraturan Menteri Koordinator Bidang Kemaritiman Dan Investasi Republik Indonesia Nomor 1 Tahun 2022	Klasifikasi Keamanan Arsip Dinamis diklasifikasikan menjadi 3 (tiga) tingkat yang terdiri atas: a. Biasa; b. Terbatas; atau c. Rahasia.	klasifikasi keamanan arsip dinamis diklasifikasikan menjadi 3 (tiga) tingkat yang terdiri atas: a. biasa; b. terbatas; atau c. rahasia.

### 2.2.2 Penghapusan Penomoran

Proses ini akan menghapus seluruh jenis penomoran. Jenis penomoran *ini* meliputi jenis penomoran huruf dan angka, baik diikuti dengan tanda titik “.” atau tanda kurung tutup “)”. Proses ini dilakukan karena penomoran dapat disalahartikan sebagai bagian dari suatu kata atau frasa (misalnya, “1.” pada “Rp1.000”). Contoh penerapan penghapusan penomoran dapat dilihat pada Tabel 6.

**Tabel 6. Contoh Penerapan Penghapusan penomoran**

Dokumen	Sebelum Pembakuan	Setelah Pembakuan
Peraturan Menteri Koordinator Bidang Kemaritiman Dan Investasi Republik Indonesia Nomor 1 Tahun 2022	klasifikasi keamanan arsip dinamis diklasifikasikan menjadi 3 (tiga) tingkat yang terdiri atas: <b>a. biasa;</b> <b>b. terbatas; atau</b> <b>c. rahasia.</b>	klasifikasi keamanan arsip dinamis diklasifikasikan menjadi 3 (tiga) tingkat yang terdiri atas: <b>biasa; terbatas; atau rahasia.</b>

### 2.2.3 Pengolahan Tanda Baca

Pengolahan tanda baca dilakukan dengan menghapus tanda baca tertentu atau mengubah tanda baca tertentu menjadi sebuah kata. Tanda baca yang dihapus adalah tanda baca yang tidak memiliki makna yang penting. Sementara itu, tanda baca yang diubah menjadi kata tertentu adalah tanda baca yang memiliki arti penting jika diubah menjadi sebuah kata.

**Tabel 7. Contoh Pengolahan Tanda Baca**



Dokumen	Sebelum Pengolahan Tanda Baca	Setelah Pengolahan Tanda Baca
Peraturan Menteri Koordinator Bidang Kemaritiman Dan Investasi Republik Indonesia Nomor 1 Tahun 2022	klasifikasi keamanan arsip dinamis diklasifikasikan menjadi 3 <b>(tiga)</b> tingkat yang terdiri <b>atas: biasa; terbatas; atau rahasia.</b>	klasifikasi keamanan arsip dinamis diklasifikasikan menjadi 3 <b>tiga</b> tingkat yang terdiri <b>atas biasa atau terbatas atau rahasia</b>

Tanda baca selain titik koma (;) dan koma (,) akan dihapus. Contoh penghapusan dapat dilihat pada Tabel 7 dengan kata yang bercetak tebal berwarna merah. Sementara itu, tanda baca perincian titik koma (;) dan koma (,) akan diolah menjadi kata *atau* atau *dan* bergantung pada kata perincian terakhirnya.

#### 2.2.4 Pembakuan Ejaan

Pembakuan ejaan ini dilakukan untuk memperbaiki ejaan kata-kata yang salah karena adanya perubahan ejaan lama menjadi ejaan baru atau adanya kesalahan penulisan yang dilakukan oleh penulis peraturan. Contoh penerapan pembakuan ejaan dapat dilihat pada Tabel 8.

**Tabel 8. Contoh Penerapan Pembakuan Ejaan**

Proses Pembakuan	Sebelum Pembakuan	Setelah Pembakuan
Perbaikan ejaan lama menjadi ejaan baru	<b>pedjabat</b> dari direktorat <b>djenderal</b> kehutanan <b>jang ditundjuk</b> oleh menteri pertanian (Peraturan Presiden Nomor 37 Tahun 1970)	<b>pejabat</b> dari direktorat <b>jenderal</b> kehutanan <b>yang ditunjuk</b> oleh menteri pertanian
Perbaikan ejaan karena kesalahan penulisan	sisa perhitungan anggaran tahun berkenaan merupakan selisih lebih dari surplus defisit <b>ditambag</b> dengan pos penerimaan pembiayaan dikurangi dengan pos pengeluaran pembiayaan daerah. (Peraturan Menteri Nomor 29 Tahun 2002)	sisa perhitungan anggaran tahun berkenaan merupakan selisih lebih dari surplus defisit <b>ditambah</b> dengan pos penerimaan pembiayaan dikurangi dengan pos pengeluaran pembiayaan daerah.

#### 2.2.5 Penghapusan Kata Umum

Proses ini akan menghapus seluruh kata-kata yang sering digunakan dalam banyak ayat dan tidak memiliki arti penting dalam pemeriksaan kontradiksi kalimat. Contoh penghapusan kata umum dapat dilihat pada tabel 9.

**Tabel 9. Contoh Penghapusan Kata Umum**

Dokumen	Sebelum Penghapusan Kata Umum	Setelah Penghapusan Kata Umum
Peraturan Menteri Koordinator Bidang Kemaritiman Dan Investasi Republik Indonesia Nomor 1 Tahun 2022	klasifikasi keamanan arsip dinamis diklasifikasikan menjadi 3 tiga tingkat <b>yang</b> terdiri atas: biasa; terbatas; atau rahasia.	klasifikasi keamanan arsip dinamis diklasifikasikan menjadi 3 tiga tingkat terdiri atas biasa atau terbatas atau rahasia

### 2.2.6 Pengolahan Kata Imbuhan

Pengolahan kata imbuhan dilakukan dengan memisahkan bentuk terikat, melakukan perluasan imbuhan, dan menghapus imbuhan. Pemisahan bentuk terikat dilakukan karena ada bentuk terikat yang memiliki bentuk antonimnya (dapat mengakibatkan dua kalimat kontradiksi), misalnya bentuk terikat “pasca” dan “pra” yang bermakna sesudah dan sebelum. Perluasan imbuhan dilakukan pada imbuhan-imbuhan kata yang perluasannya memiliki makna yang berbeda dengan kata dasarnya, sedangkan penghapusan imbuhan pada kata-kata berimbuhan lainnya dilakukan karena memiliki makna serupa dengan kata dasarnya (penghapusan imbuhan ini banyak dilakukan terhadap kata kerja dan kata benda). Contoh masing-masing pengolahan kata imbuhan dapat dilihat pada Tabel 10.

**Tabel 10. Contoh Pengolahan Kata Imbuhan**

Proses Pengolahan	Sebelum Pengolahan	Setelah Pengolahan
Pemisahan Bentuk Terikat	pemerintah dapat memberikan hibah kepada pemerintah daerah dalam rangka rehabilitasi dan rekonstruksi <b>pascabencana</b>	pemerintah dapat memberikan hibah kepada pemerintah daerah dalam rangka rehabilitasi dan rekonstruksi <b>pasca bencana</b>
Perluasan imbuhan	peserta melakukan pembayaran simpanan <b>berhak</b> memperoleh unit penyertaan investasi	peserta melakukan pembayaran simpanan <b>memiliki hak</b> memperoleh unit penyertaan investasi
Penghapusan imbuhan	peserta <b>melakukan pembayaran simpanan memiliki hak memperoleh unit penyertaan</b> investasi	peserta <b>laku bayar simpanan milik hak oleh unit serta</b> investasi

### 2.3 Data Splitting

*Data splitting* adalah proses membagi data menjadi data latih dan data uji menggunakan metode *stratified sampling*. Metode ini memastikan persentase masing-masing label di data latih dan data uji adalah kurang lebih sama. Rincian persentase distribusi label di data latih dan data uji dapat dilihat pada Tabel 11.

**Tabel 11. Hasil Pembagian Data Menggunakan *Stratified Sampling***

Label	Persentase Label di Data Latih	Persentase Label di Data Uji
Kontradiksi	25,10%	25,12%
Selaras	36,32%	36,37%
Netral	38,57%	38,50%

## 2.4 Cross Encoding

Model pembelajaran mesin hanya dapat melakukan komputasi dengan masukan berupa angka. Oleh karena itu, teks yang sudah diproses harus diubah menjadi sebuah representasi angka (tokenisasi). Proses ini dilakukan dengan memisahkan tiap kata dalam sebuah kalimat. Lalu, kata yang sudah terpisah diberikan nomor yang unik. Contoh langkah dari tokenisasi kalimat dapat dilihat pada Tabel 12.

**Tabel 12. Langkah Tokenisasi Kalimat**

Tahapan	Kalimat Pertama	Kalimat Kedua
Sebelum Tokenisasi	peserta laku bayar simpan milik hak oleh unit serta investasi	peserta laku bayar simpan tidak milik hak oleh unit serta investasi
Pemisahan Kata	[peserta, laku, bayar, simpan, milik, hak, oleh, unit, serta, investasi]	[peserta, laku, bayar, simpan, tidak, milik, hak, oleh, unit, serta, investasi]
Setelah Tokenisasi	[420, 320, 33, 441, 378, 265, 401, 477, 434, 289]	[420, 320, 33, 441, 460, 378, 265, 401, 477, 434, 289]
Cross Encoding	[420, 320, 33, 441, 378, 265, 401, 477, 434, 289, #SEP, 420, 320, 33, 441, 460, 378, 265, 401, 477, 434, 289]	

Proses *crossencoding* adalah menggabungkan hasil tokenisasi kalimat pertama dan kedua. Token #SEP digunakan sebagai pemisah antara hasil tokenisasi kalimat pertama dan kalimat kedua

## 2.5 BERT Word Embedding

BERT *word embedding* adalah teknik *word embedding* yang dapat melakukan *self-learning* menggunakan BERT. *Self-learning* ini dilakukan dengan melatih BERT menggunakan arsitektur *learning* tertentu yang menerima masukan dua kalimat berbeda dan menghitung nilai *cosine similarity* antara dua kalimat tersebut. Parameter BERT akan diperbaharui berdasarkan persamaan *loss* tertentu. Penelitian ini menggunakan dua variasi *word embedding*, yaitu Sentence BERT embedding dan SimCSE BERT Embedding. Dua variasi ini dilatih menggunakan arsitektur *learning* yang sama tetapi berbeda dari jenis persamaan *loss* yang digunakan.

### 2.5.1 Sentence BERT

Sentence BERT adalah jenis BERT *word embedding* yang dilatih dengan persamaan *mean squared error loss* (MSE):

$$MSE = \frac{1}{N} \sum_{j=1}^N (sim(u_j, v_j) - y)^2 \quad (1)$$

dengan N adalah banyaknya data,  $u_j$  dan  $v_j$  berturut-turut merupakan hasil keluaran BERT terhadap kalimat pertama dan kalimat kedua, dan  $sim(u_j, v_j)$  adalah nilai *cosine similarity* antara  $u_j$  dan  $v_j$  (Reimers & Gurevych, 2019).

### 2.5.2 SimCSE BERT

SimCSE adalah jenis BERT *word embedding* menggunakan teknik *contrastive learning* (Gao et al., 2021). Pasangan kalimat dimasukkan ke dalam arsitektur BERT dan dilatih menggunakan persamaan *contrastive learning loss*:

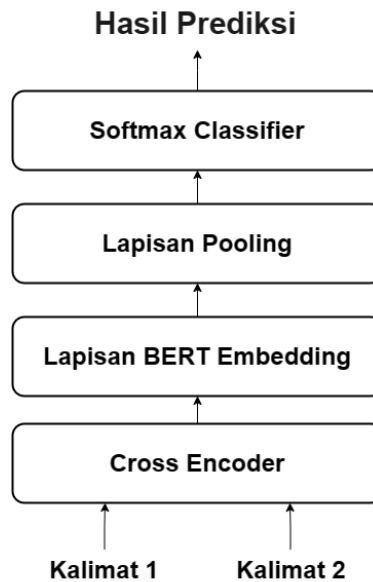
$$contrastive\ learning\ loss = -\log \frac{e^{sim(h_i, h_j^+)/\tau}}{\sum_{j=1}^N (e^{sim(h_i, h_j^+)/\tau} + e^{sim(h_i, h_j^-)/\tau})} \quad (2)$$

dengan  $\tau$  adalah parameter yang dapat dilakukan *hyperparameter tuning*.

## 2.6 Pembuatan Model

### 2.6.1 Deep Siamese Network

Deep Siamese *network* adalah arsitektur pembelajaran mesin yang menerima dua masukan berbeda untuk mencari nilai perbedaan antara dua masukan tersebut dengan persamaan tertentu (Koch, 2015).



Gambar 2. Arsitektur *Deep STS Siamese Network*

Arsitektur yang dibuat pada penelitian ini menerima satu pasang kalimat yang telah dilakukan *crossencoding*. Hasil *crossencoding* dimasukkan ke lapisan BERT *word embedding* dan lapisan *pooling*. Representasi angka yang dihasilkan oleh lapisan *pooling* dimasukkan ke lapisan *softmax* untuk melakukan klasifikasi *multiclass*.

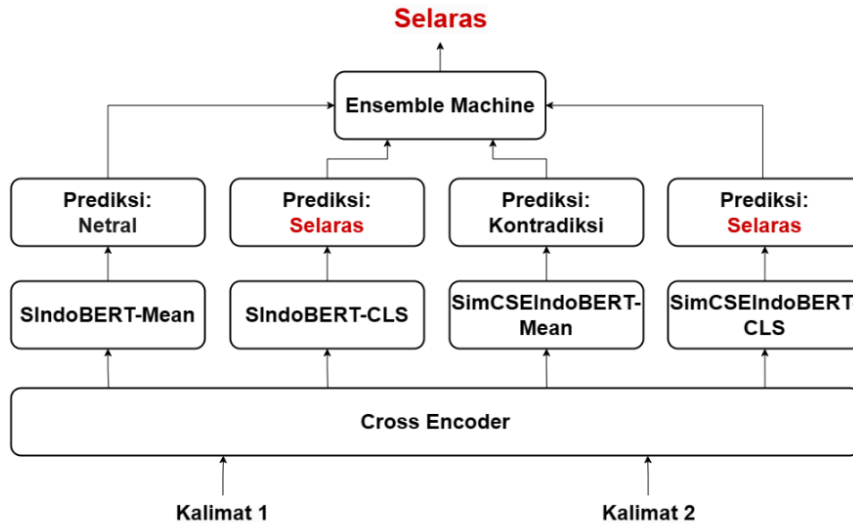
**Tabel 13. Variasi Arsitektur Deep Siamese Network**

<b>Nama Variasi</b>	<b>Lapisan <i>Embedding</i></b>	<b>Lapisan <i>Pooling</i></b>
SIndoBERT-Mean	<i>Sentence</i> IndoBERT	Mean
SIndoBERT-CLS	<i>Sentence</i> IndoBERT	CLS
SimCSEIndoBERT-Mean	SimCSE IndoBERT	Mean
SimCSEIndoBERT-CLS	SimCSE IndoBERT	CLS

Arsitektur Siamese yang digunakan pada penelitian ini memiliki dua variasi lapisan *word embedding* dan dua variasi lapisan *pooling*. Sentence IndoBERT adalah pretrained model IndoBERT yang dilatih berdasarkan teknik Sentence BERT *embedding*, sedangkan SimCSE IndoBERT adalah pretrained model IndoBERT yang dilatih berdasarkan teknik SimCSE BERT *embedding*. CLS *pooling* adalah hasil keluaran *CLS-token* dari lapisan BERT, sedangkan mean pooling adalah teknik *pooling* dengan mengambil nilai rata-rata dua token bertetangga dari keluaran lapisan BERT *embedding*. Rincian variasi arsitektur dapat dilihat pada Tabel 13.

### **2.6.2 Ensemble**

*Ensemble* adalah teknik penggabungan hasil prediksi dua atau lebih model. Teknik ini sering digunakan untuk meningkatkan performa prediksi model. Sebuah penelitian pemrosesan bahasa alami terkait permasalahan *question answering* menggunakan teknik *ensemble* terhadap model BERT (Wang et al., 2019). Hasil penelitian tersebut menunjukkan bahwa teknik *ensemble* meningkatkan nilai F1-score sebanyak 1% daripada nilai F1-score model BERT tunggal (tanpa *ensemble*).



**Gambar 3. Arsitektur Ensemble Deep STS Siamese Network**

*Ensemble* dilakukan dengan melakukan mekanisme *major voting* terhadap hasil prediksi seluruh variasi model. Ada kemungkinan label yang dihasilkan oleh empat variasi model berjumlah sama sehingga dibutuhkan satu model penentu jika hal ini terjadi. Model penentu ini akan dipilih berdasarkan performa terbaik dibandingkan dengan semua model tunggal lainnya. Ilustrasi implementasi ensemble dapat dilihat pada Gambar 2.

## 2.7 Evaluasi Model

Model dievaluasi dengan metrik-metrik, yaitu *precision*, *recall*, *F1-score*, dan akurasi. Perhitungan metrik-metrik tersebut dihitung berdasarkan nilai *true positive* (TP), *false positive* (FP), *false negative* (FN), dan *true negative* (TN) pada *confusion matrix*.

**Tabel 14. Confusion Matrix**

	<b>Aktual: True</b>	<b>Aktual: False</b>
<b>Prediksi: True</b>	<i>True Positive</i> (TP)	<i>False Positive</i> (FP)
<b>Prediksi: False</b>	<i>False Negative</i> (FN)	<i>True Negative</i> (TN)

Persamaan-persamaan metrik yang digunakan adalah:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

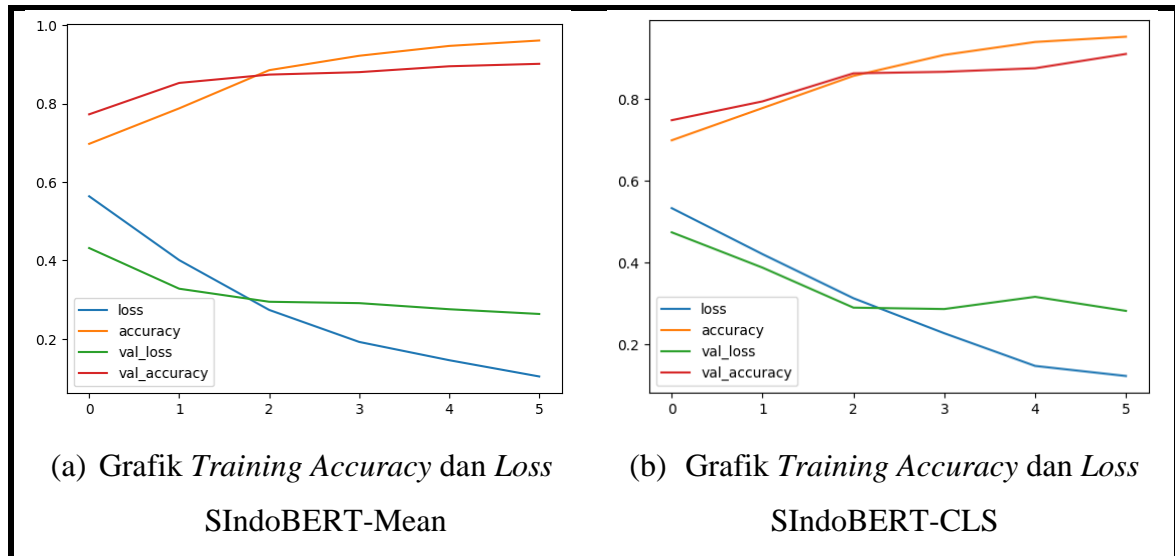
$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (6)$$

### III. PEMBAHASAN

#### 3.1 Sentence IndoBERT

Model dengan lapisan embedding Sentence IndoBERT dikombinasikan dengan lapisan *mean pooling* dan *CLS pooling*. Model ini dilatih sebanyak 6 *epoch* dengan *learning rate*  $10^{-5}$ . Grafik perubahan akurasi dan *loss* saat proses pelatihan dapat dilihat pada Gambar 3.



**Gambar 4. Grafik *Training Accuracy* dan *loss* Sentence IndoBERT**

Berdasarkan waktu pelatihan yang dibutuhkan rata-rata model memerlukan waktu 9 menit. Melihat Gambar 4 Terjadi peningkatan akurasi data latih yang signifikan pada iterasi pertama dan kedua pelatihan model SIndoBERT, tetapi tidak terjadi kenaikan signifikan setelah iterasi kedua. Kami telah mendapatkan hasil prediksi data uji ditampilkan pada tabel 15 dan 16 untuk melihat *precision*, *recall*, akurasi, dan skor F1.

**Tabel 15. *Confusion Matrix* SIndoBERT-Mean**

		Prediksi		
		Kontradiksi	Selaras	Netral
Label Sebenarnya	Kontradiksi	147	52	2
	Selaras	21	270	0
	Netral	2	2	304

Berdasarkan *confusion matrix* pada tabel 15, nilai *precision*, *recall*, *F1-score*, dan akurasi berturut-turut adalah 89,72%, 88,21%, 88,69%, dan 90,12%.

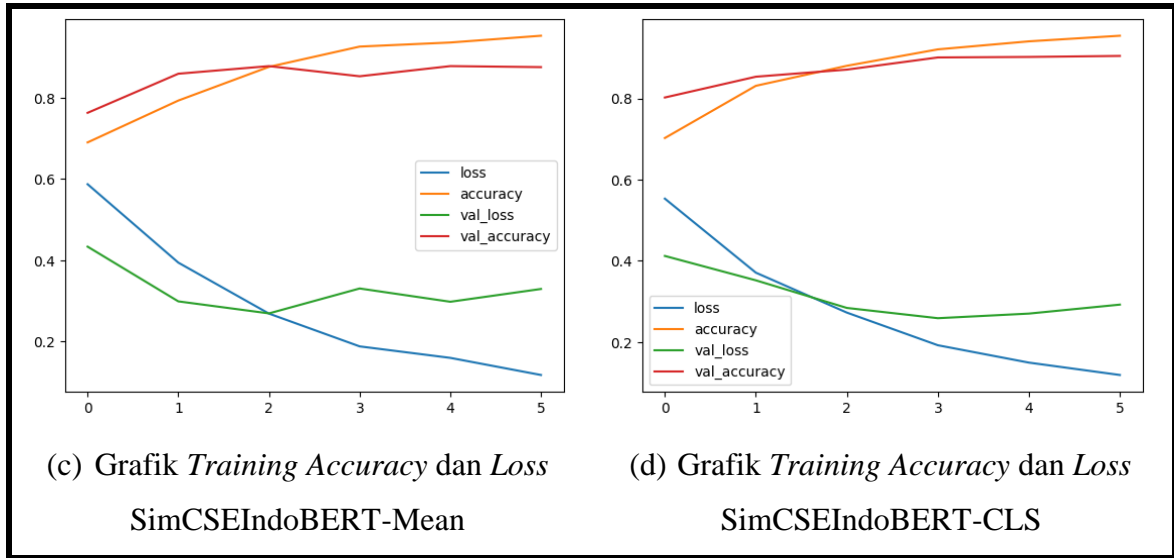
**Tabel 16. *Confusion Matrix* SIndoBERT-CLS**

		Prediksi		
		Kontradiksi	Selaras	Netral
Label Sebenarnya	Kontradiksi	161	39	1
	Selaras	27	263	1
	Netral	3	0	305

Berdasarkan *confusion matrix* pada tabel 16, nilai *precision*, *recall*, *F1-score*, dan akurasi berturut-turut adalah 90,24%, 89,83%, 90,01%, dan 91,12%.

### 3.2 SimCSE IndoBERT

Model dengan lapisan embedding SimCSE IndoBERT dikombinasikan dengan lapisan *mean pooling* dan *CLS pooling*. Model ini dilatih sebanyak 6 *epoch* dengan *learning rate*  $10^{-5}$ . Grafik perubahan akurasi dan *loss* saat proses pelatihan dapat dilihat pada Gambar 3.



**Gambar 5. Grafik Training Accuracy dan Loss SimCSE IndoBERT**

Berdasarkan waktu pelatihan yang dibutuhkan rata-rata model memerlukan waktu 9 menit. Melihat Gambar 4 Terjadi peningkatan akurasi data latih yang signifikan pada iterasi pertama dan kedua pelatihan model SimCSEIndoBERT, tetapi tidak terjadi kenaikan signifikan setelah iterasi kedua. Kami telah mendapatkan hasil prediksi data uji ditampilkan pada tabel 17 dan tabel 18 untuk melihat *precision*, *recall*, akurasi, dan skor F1.



**Tabel 17. Confusion Matrix SimCSEIndoBERT-Mean**

		Prediksi		
		Kontradiksi	Selaras	Netral
Label Sebenarnya	Kontradiksi	130	170	1
	Selaras	14	277	0
	Netral	4	3	301

Berdasarkan *confusion matrix* pada tabel 17, nilai *precision*, *recall*, *F1-score*, dan akurasi berturut-turut adalah 88,88%, 85,86%, 86,54%, dan 88,50%.

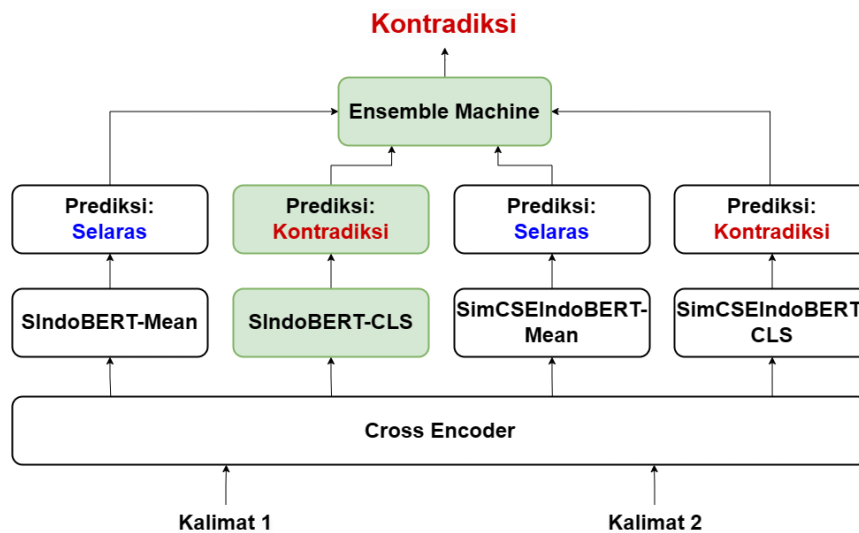
**Tabel 18. Confusion Matrix SimCSEIndoBERT-CLS**

		Prediksi		
		Kontradiksi	Selaras	Netral
Label Sebenarnya	Kontradiksi	167	33	1
	Selaras	43	248	0
	Netral	4	1	303

Berdasarkan *confusion matrix* pada tabel 18, nilai *precision*, *recall*, *F1-score*, dan akurasi berturut-turut adalah 88,55%, 88,89%, 88,69%, dan 89,75%.

### 3.3 Ensemble

SIndoBERT-CLS menghasilkan performa paling baik dibandingkan model tunggal lainnya. Oleh karena itu, SIndoBERT-CLS ini dijadikan sebagai penentu hasil prediksi *ensemble* jika terjadi kesamaan jumlah label hasil prediksi variasi model tunggal. Ilustrasi prediksi *ensemble* ketika terjadi kesamaan jumlah label dapat dilihat pada Gambar 6.



**Gambar 6. Prediksi Model Ensemble Ketika Terjadi Kesamaan Jumlah Label**

Model *ensemble* memiliki performa lebih baik daripada seluruh model tunggal berdasarkan seluruh metrik evaluasi yang digunakan. Hal ini disebabkan oleh mekanisme

*major voting* yang dilakukan pada model *ensemble* sehingga kesalahan klasifikasi satu model dapat ditutupi oleh hasil klasifikasi model tunggal lainnya.

**Tabel 19. Hasil Evaluasi Model**

Arsitektur Model	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	Akurasi
SIndoBERT-Mean	89,72%	88,21%	88,69%	90,12%
SIndoBERT-CLS	90,24%	89,83%	90,01%	91,12%
SimCSEIndoBERT-Mean	88,88%	85,86%	86,54%	88,50%
SimCSEIndoBERT-CLS	88,55%	88,89%	88,69%	89,75%
Ensemble	<b>91,09%</b>	<b>90,41%</b>	<b>90,68%</b>	<b>91,75%</b>

Secara keseluruhan hasil evaluasi model terhadap Tabel 19 menunjukkan bahwa variasi arsitektur word embedding dan pooling layer memberikan pengaruh lebih baik terhadap kinerja model dalam mengidentifikasi antar kelasnya. Arsitektur model Ensemble meraih performa lebih baik daripada penggunaan model tunggal dengan akurasi 91,75% dan *F1-score* 90,68%. Contoh evaluasi model prediksi berhasil dapat dilihat pada Tabel 20 dan prediksi gagal pada Tabel 21.

**Tabel 20. Contoh Keberhasilan Prediksi Model Ensemble**

Kalimat Pertama	Kalimat Kedua	Label	Prediksi
Defisit anggaran sebagaimana dimaksud pada Ayat (5) terjadi apabila Anggaran Pendapatan Daerah <b>lebih kecil</b> dari Anggaran Belanja Daerah.	Defisit anggaran sebagaimana dimaksud pada Ayat (5) terjadi apabila Anggaran Pendapatan Daerah <b>lebih besar</b> dari Anggaran Belanja Daerah.	Kontradiksi	Kontradiksi

Keberhasilan model dalam memprediksi kalimat pertama dengan kedua adalah keberhasilan prediksi model dalam mengidentifikasi antonim yaitu kata atau frasa yang memiliki makna berlawanan atau berkebalikan.

**Tabel 21. Contoh Kegagalan Prediksi Model**

Kalimat Pertama	Kalimat Kedua	Label	Prediksi
Defisit anggaran sebagaimana dimaksud pada Ayat (5) terjadi apabila <b>Anggaran Pendapatan Daerah lebih kecil dari Anggaran Belanja Daerah.</b>	Defisit anggaran sebagaimana dimaksud pada Ayat (5) terjadi apabila <b>Anggaran Belanja Daerah lebih besar dari Anggaran Pendapatan Daerah.</b>	Selaras	Kontradiksi
Bank dan Pihak Terafiliasi wajib <b>merahasiakan</b> keterangan	Bank dan Pihak Terafiliasi wajib <b>mengumumkan ke publik</b> keterangan	Kontradiksi	Selaras

Kalimat Pertama	Kalimat Kedua	Label	Prediksi
mengenai Nasabah Penyimpan dan Simpanannya serta Nasabah Investor dan Investasinya.	mengenai Nasabah Penyimpan dan Simpanannya serta Nasabah Investor dan Investasinya.		

Berdasarkan contoh pertama kegagalan prediksi model pada Tabel 21, model terlalu fokus dalam mengenali bahwa kata “besar” adalah antonim dari kata “kecil” sehingga model memprediksi dua kalimat tersebut berkontradiksi. Padahal, kata antonim tersebut juga diikuti dengan pertukaran posisi dua objek perbandingan sehingga dua kalimat tersebut seharusnya masih selaras. Contoh kedua juga merupakan kegagalan prediksi model belum dapat mengenali bahwa merahasiakan memiliki makna berlawanan dengan mengumumkan ke publik. Dua contoh kegagalan tersebut dapat diatasi dengan menambah *dataset* baru dengan variasi yang serupa agar model dapat mempelajari.

## IV. PENUTUP

### 4.1 Kesimpulan

Dalam rangka membantu pemerintah dalam pembuatan sistem identifikasi otomatis pemeriksaan keselarasan dan kontradiksi peraturan baru terhadap peraturan yang sudah ada, penelitian ini membuat model dengan dasar arsitektur *deep Siamese network*. Model Siamese dibuat dengan kombinasi lapisan *embedding* dan *pooling* yang bervariasi, yaitu SIndoBERT-Mean, SIndoBERT-CLS, SimCSEIndoBERT-Mean, dan SimCSEIndoBERT-CLS. Kombinasi model tunggal terbaik adalah SIndoBERT-CLS. Kombinasi tersebut mencapai tingkat akurasi 91,12% dan *F1-score* sebesar 90,01%. Selain itu, performa model ini dapat ditingkatkan dengan menggabungkan keempat model tersebut (Ensemble Deep STS Siamese Network) yang menghasilkan akurasi 91,75% dan *F1-score* sebesar 90,68%. Model ensemble yang dibuat berhasil meningkatkan performa model karena kesalahan klasifikasi satu model tunggal dapat ditutupi oleh hasil prediksi model tunggal lainnya.

### 4.2 Saran

Performa model *Ensemble Deep STS Siamese Network* sudah menunjukkan performa yang baik dalam mengidentifikasi kontradiksi dan keselarasan. Namun, penelitian ini masih dapat dikembangkan lagi dengan:

1. Menambahkan variasi dataset peraturan yang selaras dan kontradiksi sehingga model dapat mengenali kalimat yang lebih variatif;
2. Bekerja sama dengan lembaga pemerintahan untuk mengakses dokumen-dokumen peraturan yang lebih banyak; dan

3. Menggunakan *Large Language Model* untuk mengetahui alasan dan pendapat dalam mengidentifikasi keselarasan dan kontradiksi dari kedua aturan yang dibandingkan. Oleh karena itu, perlunya dataset *question and answering* (QA) tentang argumentasi pakar hukum terhadap seluruh peraturan yang ada.

## REFERENSI

- Gao, T., Yao, X., & Chen, D. (2021). SimCSE: Simple Contrastive Learning of Sentence Embeddings. *CoRR*, *abs/2104.08821*. <https://arxiv.org/abs/2104.08821>
- Koch, G. R. (2015). *Siamese Neural Networks for One-Shot Image Recognition*. <https://api.semanticscholar.org/CorpusID:13874643>
- Nurhardianto, F. (2015). Sistem Hukum dan Posisi Hukum Indonesia. *Jurnal Teropong Aspirasi Politik Islam*, *11*(1), 35–45.
- Nurjaman, J., Ilyas, R., & Kasyidi, F. (2020). Pengukuran Kesamaan Semantik Pasangan Kalimat Sitasi Menggunakan Convolutional Neural Network. *Prosiding Industrial Research Workshop and National Seminar*.
- Rajagukguk, R. C., & Leylia Khodra, M. (2018). Interpretable Semantic Textual Similarity for Indonesian Sentence. *2018 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA)*, 147–152. <https://doi.org/10.1109/ICAICTA.2018.8541297>
- Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *CoRR*, *abs/1908.10084*. <http://arxiv.org/abs/1908.10084>
- Sumodiningrat, A., Helmi, A., & Rifat, T. (2022). Desain Lembaga Khusus Bidang Legislasi dan Evaluasi Peraturan Perundang-undangan Partisipatif. *Jurnal Legislasi Indonesia*, *19*(3), 380–395.
- Usino, W., Prabuwo, A. S., Allehaibi, K. H. S., Bramantoro, A., A, H., & Amaldi, W. (2019). Document Similarity Detection using K-Means and Cosine Distance. *International Journal of Advanced Computer Science and Applications*, *10*(2). <https://doi.org/10.14569/IJACSA.2019.0100222>
- Wang, W., Bi, B., Yan, M., Wu, C., Bao, Z., Xia, J., Peng, L., & Si, L. (2019). *StructBERT: Incorporating Language Structures into Pre-training for Deep Language Understanding*.