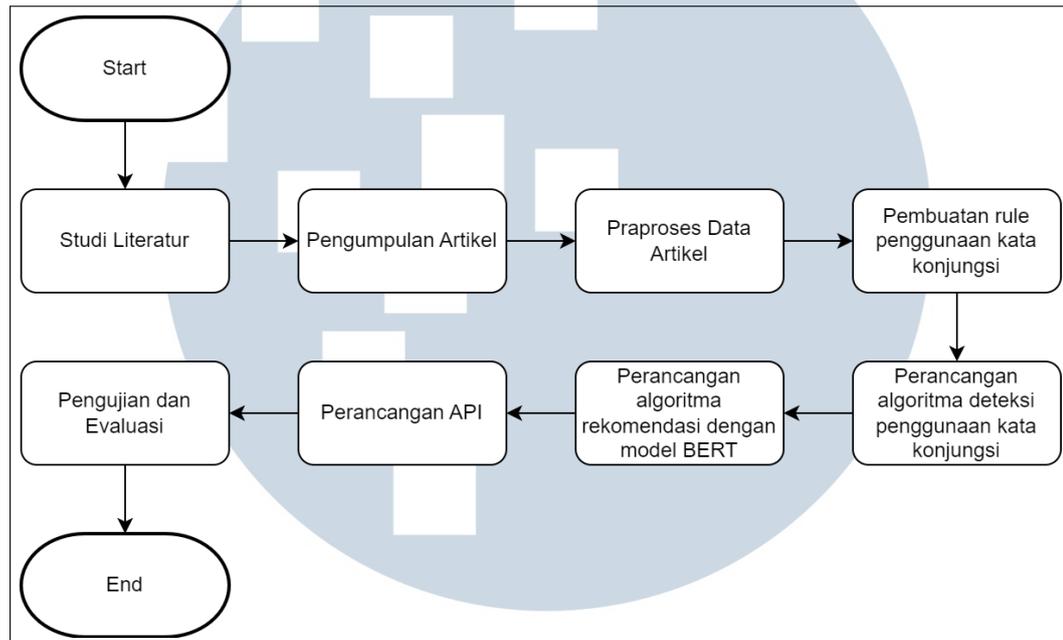


BAB 3 METODOLOGI PENELITIAN

Metodologi penelitian untuk penelitian ini dapat dilihat pada 3.1.



Gambar 3.1. Diagram alur metodologi penelitian

3.1 Studi Literatur

Pada tahap studi literatur, dilakukan penelitian teori yang dapat digunakan sebagai dasar dalam pengembangan sistem deteksi kesalahan penggunaan kata konjungsi menggunakan *Semantic Analysis* dan sistem saran perbaikan untuk penggunaan kata konjungsi yang salah menggunakan model BERT.

3.2 Pengumpulan dan Praproses Data Artikel

Pada tahap pengumpulan artikel, dilakukan pengumpulan data artikel berita dari portal berita Tribunnews.com. Data yang sudah dikumpulkan ini digunakan untuk melakukan perhitungan tingkat akurasi dari sistem deteksi penggunaan kata konjungsi yang sudah dibuat. Artikel yang sudah dikumpulkan kemudian dilakukan praproses data untuk mengganti data yang berpotensi dapat menyebabkan error pada sistem. Praproses yang dilakukan adalah mengubah karakter non-ASCII

menjadi karakter ASCII terdekat. Contohnya adalah karakter petik dua buka dan tutup dengan kode *unicode* "U+201C" dan "U+201D" menjadi karakter petik dua dengan kode *unicode* "U+0022". Hal ini juga berlaku untuk huruf non-Latin, huruf non-Latin juga diubah menjadi huruf Latin yang paling mendekati.

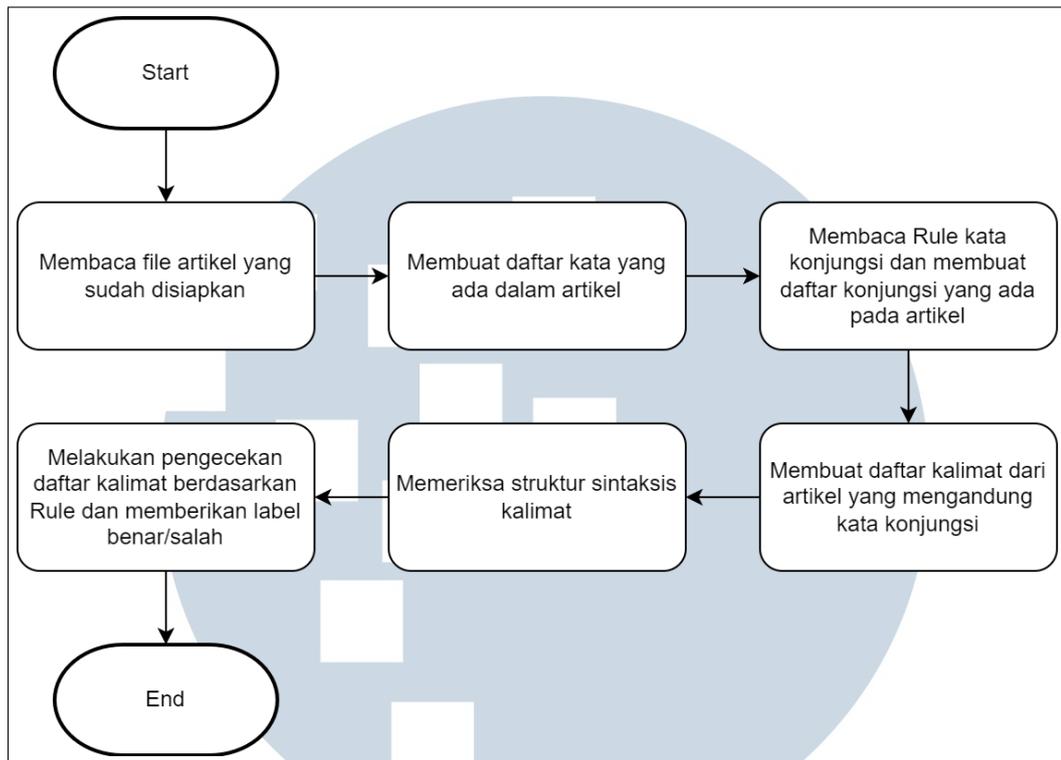
3.3 Pembuatan *rule* penggunaan kata konjungsi

Pada tahap pembuatan *rule* penggunaan kata konjungsi, dilakukan diskusi dengan Ibu Niknik untuk menentukan daftar kata konjungsi beserta kondisi yang harus dipenuhi dalam penggunaan kata konjungsi. *Rule* ini kemudian disesuaikan untuk meningkatkan efektivitas saat digunakan dalam proses pemeriksaan penggunaan kata konjungsi. *Rule* yang sudah disesuaikan kembali didiskusikan dengan Ibu Niknik untuk mencegah terjadinya atau memperbaiki kesalahan saat melakukan penyesuaian. Hasil pembuatan *rule* yang pertama dapat dilihat pada Gambar 3.2 yang berupa daftar aturan penggunaan kata konjungsi. Dapat dilihat terdapat beberapa aturan yang memiliki *CHILD_RULE* yang merupakan aturan tambahan yang harus dipenuhi agar aturan yang memiliki *CHILD_RULE* tersebut dapat terpenuhi.

ID	RULE	CHILD_RULE	RULE_TYPE	FORMULA
0	Intra kalimat (tengah)		regex	<conj>
1	Intra kalimat (awal)	11	regex	^<conj>
2	Antar kalimat	5,10	regex	^<conj>
3	Kalimat setara	8	needChild	
4	Kalimat tidak setara	9	needChild	
5	Jika awal kalimat, gunakan koma	11	regexC	^<conj> ,
6	Kata konjungsi lebih dari 1 dan tidak bersebelahan	7	regex	<conj1> \S+ <conj2> \S
7	Menggabungkan 3 atau lebih unsur menggunakan koma	8	regex	(\w+ ,)+<conj> \w+
8	Koma sebelum konjungsi (tetapi, melainkan, sedangkan, kecuali)		regex	, <conj>
9	Jika anak kalimat diawal, pakai koma, jika dibelakang, tidak pakai koma		regex	(^<conj>-(?! ,)+,) (^[,]+⟩
10	Tidak di awal paragraf		expression	row['ID'] > 0
11	Diawali huruf kapital		regexC	<conj>

Gambar 3.2. Daftar aturan penggunaan kata konjungsi

Pada Gambar 3.3 merupakan daftar kata konjungsi beserta aturan yang harus dipenuhi agar penggunaan kata konjungsi tersebut dapat dianggap benar.

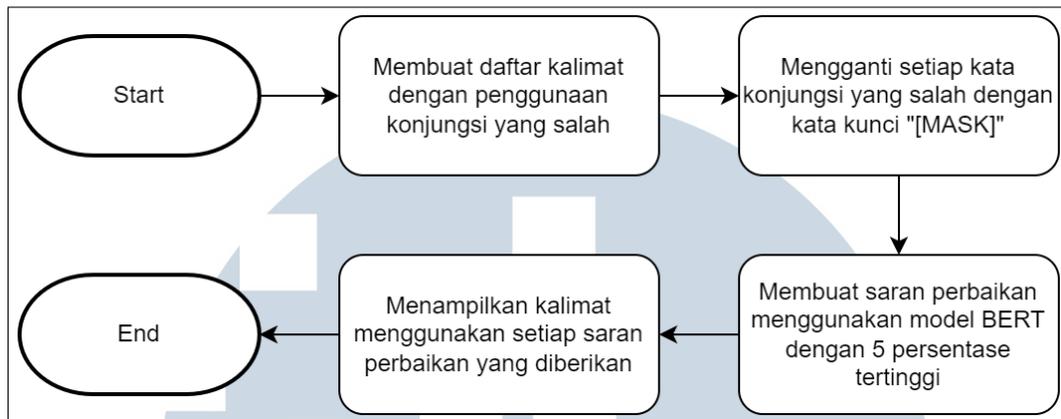


Gambar 3.4. Diagram alur algoritma deteksi kesalahan penggunaan kata konjungsi

3.5 Implementasi model BERT untuk saran perbaikan kata konjungsi

Model BERT diimplementasikan untuk memberikan saran perbaikan kata konjungsi dengan memasukkan kalimat dengan kata konjungsi yang salah dan mengganti kata konjungsi tersebut dengan kata kunci "[MASK]". Lalu memasukkan kalimat tersebut ke dalam model untuk menghasilkan saran perbaikan kata. Model BERT yang digunakan adalah model yang sudah dilatih menggunakan dataset dari situs Wikipedia dalam Bahasa Indonesia sebesar 522 MB. Alur proses algoritma saran perbaikan dapat dilihat pada Gambar 3.5

UNIVERSITAS
MULTIMEDIA
NUSANTARA



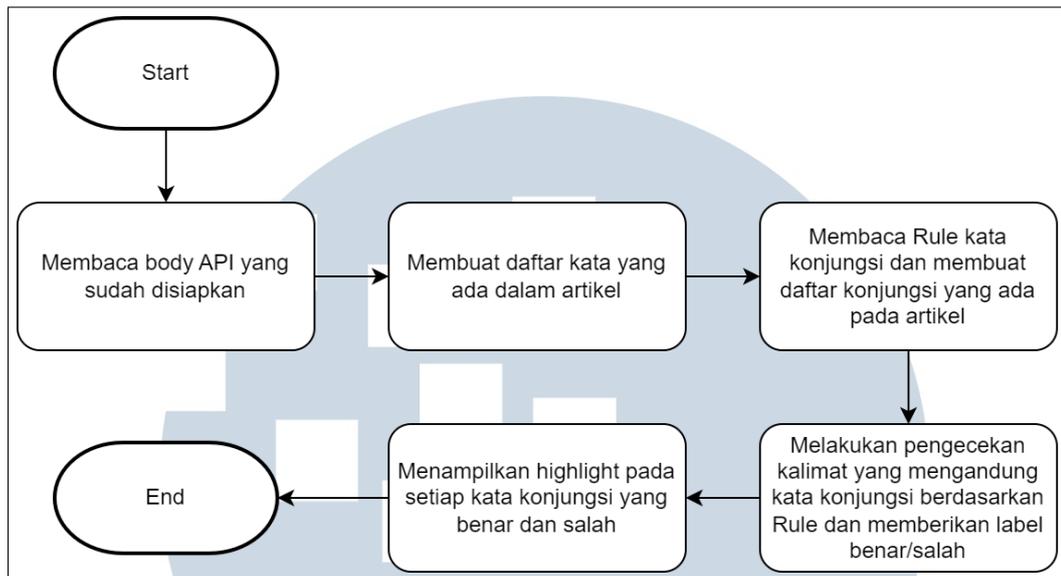
Gambar 3.5. Diagram alur algoritma saran perbaikan kata konjungsi

3.6 Perancangan API

Pada tahap perancangan API, dilakukan implementasi program deteksi kesalahan penggunaan kata konjungsi dengan membuat *Application Programming Interface* (API) menggunakan *framework* Flask dengan bahasa pemrograman Python. Implementasi program dalam bentuk API dapat mempermudah implementasi program dalam bentuk media lainnya, seperti *website*, *mobile application*, dan lainnya.

Program Flask yang sudah di *deploy* dapat diakses melalui *endpoint* `http://{hostname}:{port}/api/v1/utapis_cek_penggunaan_konjungsi` dengan metode *POST request*. Untuk parameter yang dikirimkan menggunakan *form-data* berisi `'article={article}'` dan akan mengembalikan artikel beserta *highlight* pada kata konjungsi yang terdeteksi. Jika penggunaan kata konjungsi benar, maka *highlight* yang diberikan adalah warna hijau, sedangkan untuk penggunaan kata konjungsi salah akan diberikan *highlight* berwarna merah yang dapat ditekan untuk memunculkan daftar saran kata perbaikan. Alur proses API dapat dilihat pada Gambar 3.6.

UNIVERSITAS
MULTIMEDIA
NUSANTARA



Gambar 3.6. Diagram alur API deteksi penggunaan kata konjungsi

3.7 Pengujian dan Evaluasi

Pada tahap pengujian dan evaluasi, dilakukan perhitungan akurasi menggunakan *Confusion Matrix* dengan melakukan perbandingan deteksi kesalahan penggunaan kata konjungsi pada program dengan label dari artikel yang sudah dikumpulkan dari media berita Tribunnews.com dan sudah dilakukan pemeriksaan oleh tim Jurnalistik.

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA