Pengisian poin C sampai dengan poin H mengikuti template berikut dan tidak dibatasi jumlah kata atau halaman namun disarankan seringkas mungkin. Dilarang menghapus/memodifikasi template ataupun menghapus penjelasan di setiap poin.

C. HASIL PELAKSANAAN PENELITIAN: Tuliskan secara ringkas hasil pelaksanaan penelitian yang telah dicapai sesuai tahun pelaksanaan penelitian. Penyajian meliputi data, hasil analisis, dan capaian luaran (wajib dan atau tambahan). Seluruh hasil atau capaian yang dilaporkan harus berkaitan dengan tahapan pelaksanaan penelitian sebagaimana direncanakan pada proposal. Penyajian data dapat berupa gambar, tabel, grafik, dan sejenisnya, serta analisis didukung dengan sumber pustaka primer yang relevan dan terkini.

Berdasarkan rencana penelitian yang telah disampaikan pada proposal penelitian, maka pada bagian ini akan menyampaikan capaian dari penelitian yang telah dilakukan:

1. Hasil Pengumpulan Data:

Pengembangan model Integrasi Deep Learning dan Knowledge Graph untuk Ekstraksi Entitas Medis, dibutuhkan sekumpulan dara dari rekam medis elektronik. Data dikumpulkan dari Rumah Sakit Umum Pusat Nasional dr. Cipto Mangunkusumo. Hasil dari pengumpulan data ini dijadikan dataset dalam penilitan yang dilakukan. Data yang dikumpulkan berfokus pada formulir pasien dewasa rawat inap, karena formulir tersebut yang sudah diimplementasikan di Rumah Sakit. Atribut pada formulir pengkajian awal pasien dewasa rawat inap dipilih yang menjadi mandatory diantaranya keluhan utama, riwayat penyakit sekarang, riwayat pengobatan, riwayat penyakit dalam keluarga, pemeriksaan fisik dan daftar masalah. Jumlah data yang dapat dikumpulkan sebanyak 41404 data, Contoh data yang dimpulkan seperti pada Tabel 1.

Tabel 1. Contoh data yang sudah terkumpul

ORDER_ID	OBJ_ID	OBS_DTTM	OBJ_NM	VALUE_LONG
135766047	30389	2022-01-01 06:58:25	KELUHAN UTAMA	Pasien Perempuan berusia 43 tahun datang dari IGD dengan keluhan muntah berulang sejak 3 bulan smrs.
135766047	30391	2022-01-01 06:58:25	RIWAYAT PENYAKIT SEKARANG	Sebelumnya pasien mengeluh nyeri ulu hati berulang, pasien minum obat lambung, namun semakin lama, pasien merasa sulit menelan.
135766047	30394	2022-01-01 06:58:25	RIWAYAT PENYAKIT DALAM KELUARGA	Riwayat HT, DM, keganasan pada keluarga disangkal.
...	...	...	...	...
135766047	42328	2022-01-01 06:58:25	DAFTAR MASALAH ATAU DIAGNOSIS	Riwayat Operasi Kista Endometrium (2009 dan 2015)

Pada Tabel 1 terdiri 5 kolom/atribut yaitu ORDER_Id, OBJ_ID, OBS_DTTM, OBJ_NM dan VALUE_LONG.

VALUE_LONG merupakan merupakan atribut yang berisi riwayat penyakit sekarang, riwayat pengobatan, riwayat penyakit dalam keluarga, keluhan utama, pemeriksaan fisik, dan daftar masalah atau diagnosis dalam bentuk naratif. Atribut VALUE_LONG ini yang akan digunakan sebagai Ekstraksi Entitas Medis.

2. Prapemrosesan Data dan Ektrasi Fitur:

Prapemrosesan data yang dilakukan adalah stopword, lammatizer, dan tokenizing. Data yang dilakukan prapemrosesan data awala adalah kolom/atribut VALUE LONG (lihat Tabel1). Berikut ini dijelaskan proses yang dilakukan:

Stopword: Langkah awal proses data preprocessing pada penelitian ini adalah proses stopword, yaitu menghilangkan kata penghubung dan kata yang terdaftar pada stopword library NLP-ID (1).

Lammatizer: langkah berikutnya pada prapemrosesam data adalah lammatizer, yaitu proses untuk mendapatkan kata dasar, kata menjadi huruf kecil dan menghilangkan angka dari sebuah kalimat.

Tokenizing: langkah terakhir dalam pramerosesan data adalah menjadikan kalimat menjadi bagian-bagian yang lebih kecil yang disebut token, token ini berupa kata, tanda baca, angka, tanggal, dan lain sebagainya.

Tabel 2. Contoh hasil prapemrosesan data

VALUE LONG	HASIL_STOPWORDS	HASIL_LEMMATIZER	HASIL_TOKENIZING
Pasien konsulan TS IGD dengan keluhan nyeri perut kanan bawah sejak 2 hari yang lalu	pasien konsulan TS IGD keluhan nyeri perut kanan	pasien konsul ts igd keluh nyeri perut kanan	pasien
			konsul
			ts
			igd
			keluh
			nyeri
			perut
			kanan
…	…	…	…
Pasien pro egd ligasi ts Gastro	Pasien pro egd ligasi ts Gastro	pasien pro egd ligasi ts gastro	pasien
			pro
			egd
			ligasi
			ts
			gastro

Berdasarkan data pada Tabel 2, maka data tersebut digunakan untuk proses selanjutnya adalah melakukan Anotasi dan Modeling. Anotasi adalah proses pemberian label pada data teks sehingga algoritme pembelajaran mesin dapat memproses dan memahaminya (2). Notasi teks adalah proses pelabelan dan penandaan elemen dalam teks agar model AI dan Pemrosesan Bahasa Alami (NLP) dapat memahami, menafsirkan, dan memproses bahasa manusia (3). Proses ini melibatkan penambahan metadata (informasi tentang data) ke teks, yang membantu model mengenali entitas, sentimen, maksud, hubungan, dan lainnya. Pada penelitian ini akan anotasi dengan metode Anotasi Entitas (NER – Pengenalan Entitas Bernama).

3. Anotasi Entitas

Anotasi entitas dilakukan dengan pendekatan pakar (dokter), berkolaborasi dengan dokter untuk penentuan label entitas (diagnosa, gejala, prosedur) dan pembuatan panduan anotasi.

Rancangan dasar untuk mendefinisikan kelas entitas pada domain rekam medis elektronik di penelitian ini dapat dilihat pada Tabel 3.

Tabel 3. Entitas Bernama

NO.	KELAS ENTITAS	BIO TAGGING	KETERANGAN
1.	Symptoms	B-sym dan I-sym	menunjukkan gejala atau tanda-tanda
2.	Body	B-bod dan I-bod	menunjukkan bagian tubuh
3.	Treatment	B-tre dan I-tre	menunjukkan pemeriksaan
4.	Nursing	B-nur dan I-nur	menunjukkan tindakan perawatan
5.	Disease	B-dis dan I-dis	menunjukkan penyakit atau diagnose
6.	Abbreviation	B-abb danI-abb	menunjukkan singkatan
7.	Other	O	menunjukkan kata yang tidak memiliki makna

Anotasi manual entitas bernama dilakukan oleh beberapa pakar, yaitu dokter spesialis ilmu penyakit dalam sebagai evaluator rekam medis elektronik, dokter spesialis ilmu kejiwaan sebagai ketua panitia rekam medis, dokter umum sebagai manager layanan medik, perawat klinis dan perekam medis sebagai trainer rekam medis elektronik. Beberapa pakar terlibat bertujuan untuk memperoleh data entitas yang lebih baik dan meningkatkan kualitas dari data.

Anotasi manual entitas bernama terlihat cotohnya pada Tabel 4, yang digunakan untuk mengetahui kelas kata dari masing-masing kata, di mana kelas kata ini akan digunakan untuk menentukan kata-kata yang memiliki entitas. Penelitian ini menggunakan format BIO merupakan singkatan dari Beginning (B), Inside (I), dan Outside (O). Anotasi manual entitas Bernama ini dapat mengklasifikasikan kata ke dalam 7 Entity Class dan 13 Tagging.

Tabel 4. contoh tagging entitas bernama

TOKENISASI	ANOTASI MANUAL ENTITAS BERNAMA
pasien konsul ts igd keluh nyeri perut kanan	pasien, 'O' konsul, 'O' ts, 'B-abb' igd, 'B-abb’ keluh, 'O' nyeri, 'B-sym' perut, 'B-bod' kanan, 'I-bod'
...	...
Pasien pro egd ligase ts gastro	pasien, ‘O’ pro, ‘O’ egd,’B-abb’ ligasi, ‘B-tre’ ts,’B-abb’ gastro, ‘O’

Berikut ini contoh dataset setelah dilakukan proses tagging. Data yang digunakan sebanyak 41404 data, yang terdiri dari sentence dan Tag. Sentence merupakan hasil dari prapemrosesan data dan Tak merupakan hasil dari proses Anotasi Entitas (lihat Tabel 5.)

Tabel 5. Contoh kumpulan data untuk modeling

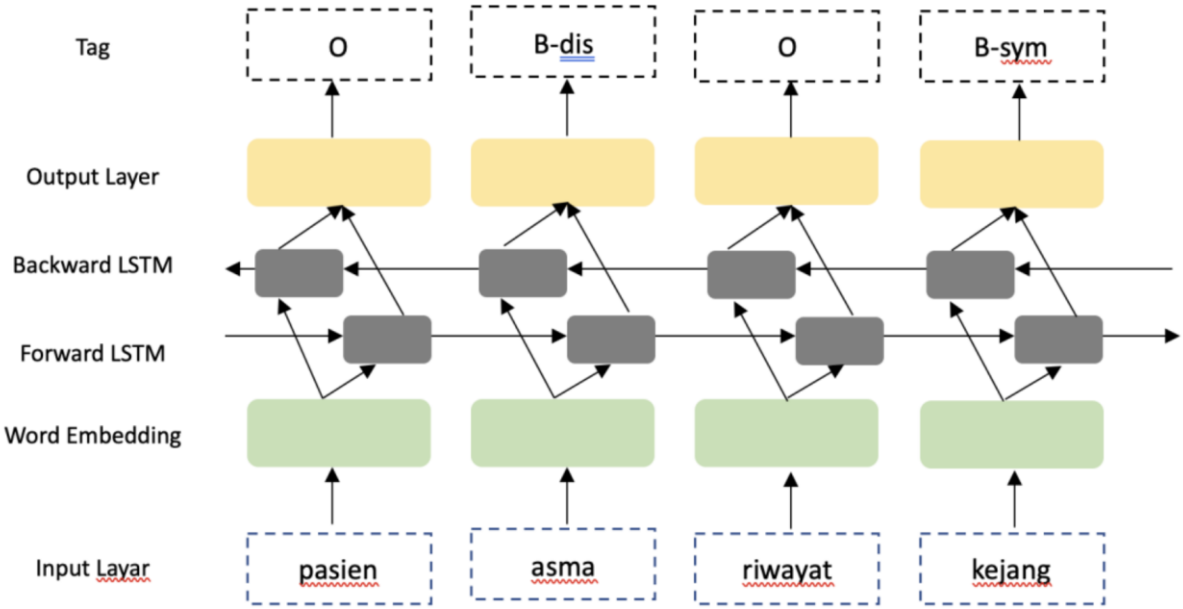
No	Sentence	Tag
1	pasien konsul ts ipd keluh nyeri perut kanan	['O', 'O', 'B-abb', 'B-abb', 'O', 'B-sym', 'B-bod', 'I-bod']
2	pasien rawat ts ipd riwayat demam dengue pasien rawat igd keluh demam turun beri oabat turun panas riwayat mimisan sangkal pasien bitnik merah sangkal bab warna hitam campur darah sangkal periksa laboratorium dapat turun trombosit pasien pasien diagnosis demam dengue pasien rawat gedung pasien keluh nyeri perut kanan rasa terus nyeri berat makan minum nyeri rasa jalar nyeri pengaruh ubah posisi mual muntah sangkal buang air keluh pasien demam mual muntah bab bantu yall volume konsistensi padatlunak warna kuning coklat	['O', 'O', 'B-abb', 'B-abb', 'O', 'B-sym', 'I-sym', 'O', 'O', 'B-abb', 'O', 'B-sym', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-abb', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-sym', 'I-sym', 'O', 'O', 'O', 'O', 'O', 'B-sym', 'B-bod', 'I-bod', 'O', 'O', 'B-sym', 'O', 'O', 'O', 'B-sym', 'O', 'O', 'B-sym', 'O', 'O', 'O', 'B-sym', 'B-sym', 'O', 'O', 'O', 'O', 'O', 'B-sym', 'B-sym', 'B-sym', 'B-abb', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O']
3	susp appendisitis perforasi dd tuba ovarial abses dextrahidroureter hidronefrosis kanan	['B-abb', 'B-dis', 'O', 'B-abb', 'O', 'O', 'B-sym', 'O', 'B-dis', 'I-bod']
4	kes compos mentis td mmhg hr min rr min mata konjungtiva anemis sklera ikterik jantung reguler murmur gallop paru vesikuler lapang paru rhonki wheezing abdomen cembung bising usus kesan normal supel nyeri tekan abdomen kanan regio kanan raba padat ukur xxcm batas terfiksir muka nyeri tekan kulit sekitar suhu raba kulit timpani ekstremitas hangat crt detik rt tsa ampula rekti kolaps mukosa licin raba sarung tangan feses darah	['O', 'B-sym', 'I-sym', 'B-abb', 'B-abb', 'B-abb', 'B-abb', 'B-abb', 'B-abb', 'B-bod', 'O', 'B-dis', 'B-sym', 'O', 'B-bod', 'O', 'B-sym', 'B-sym', 'B-bod', 'I-sym', 'O', 'B-bod', 'B-sym', 'B-sym', 'I-tre', 'O', 'O', 'B-bod', 'O', 'O', 'O', 'B-sym', 'b-nur', 'I-tre', 'I-bod', 'O', 'I-bod', 'O', 'O', 'O', 'B-abb', 'O', 'O', 'B-bod', 'B-sym', 'b-nur', 'B-bod', 'O', 'O', 'O', 'B-bod', 'O', 'B-sym', 'O', 'B-abb', 'O', 'B-abb', 'B-abb', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-bod', 'O', 'O']
5	pasien perempuan usia igd keluh muntah ulang smrs	['O', 'O', 'O', 'B-abb', 'O', 'B-sym', 'O', 'B-abb']
6	pasien perempuan usia igd keluh muntah ulang smrs smrs mual muntah makan keluh habis makan bubur muntah makan nasi telan muntah nafsu makan kurang pasien makan sendok keluh bab darah bab hitam sangkal muntah darah sangkal smrs nafsu makan kurang pasien minum susu telan makan padat turun bb kg pasien keluh nyeri ulu hati ulang pasien minum obat lambung pasien sulit telan keluh demam batuk sesak napas nyeri dada sangkal pasien awalanya obat rs sari asih serang smrs pasien rawat muntah ulang dapat usus buntu operasi usus buntu pasien rujuk rs krakatau timbang endoskopi pasien obat rawat jalan kgeh krakatau medika jalan egd desember rujuk rscm smrs pasien diet susu minum susu air putih pasien muntah rpd riwayat ht dm ganas sakit jantung ginjal tb sangkal rpk riwayat keluh keluarga sangkal riwayat ht dm ganas keluarga sangkal sosial pasien nikah milik anak pasien milik biasa rutin minum jamu usia an riwayat rokok minum alkohol obat herbal sangkal igd pasien diassess syok hipovolemik loading cair pasien egd pasang ngt pasien tpn	['O', 'O', 'O', 'B-abb', 'O', 'B-sym', 'O', 'B-abb', 'B-abb', 'B-sym', 'B-sym', 'O', 'O', 'O', 'O', 'O', 'B-sym', 'O', 'O', 'b-nur', 'B-sym', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-abb', 'O', 'B-abb', 'O', 'O', 'B-sym', 'O', 'O', 'B-abb', 'O', 'O', 'O', 'O', 'O', 'O', 'b-nur', 'O', 'O', 'O', 'B-abb', 'B-abb', 'O', 'O', 'B-sym', 'I-sym', 'B-bod', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'b-nur', 'O', 'B-sym', 'B-sym', 'B-sym', 'O', 'B-sym', 'B-bod', 'O', 'O', 'O', 'O', 'B-abb', 'O', 'O', 'O', 'B-abb', 'O', 'O', 'B-sym', 'O', 'O', 'B-bod', 'O', 'B-tre', 'B-bod', 'O', 'O', 'O', 'B-abb', 'O', 'O', 'B-tre', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-abb', 'O', 'O', 'B-abb', 'B-abb', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-sym', 'B-abb', 'O', 'B-abb', 'B-abb', 'O', 'O', 'B-bod', 'B-bod', 'B-abb', 'O', 'B-abb', 'O', 'O', 'O', 'O', 'O', 'B-abb', 'B-abb', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-abb', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-abb', 'O', 'B-abb', 'O', 'O']

No	Sentence	Tag
7	rpk riwayat keluh keluarga sangkal riwayat ht dm ganas keluarga sangkal	['B-abb', 'O', 'O', 'O', 'O', 'O', 'B-abb', 'B-abb', 'O', 'O', 'O']
8	gastric outlet obstruction massa korpus antrum post egd guided ngt esofagitis erosiva la class bc malnutrisi risiko very high risk refeeding syndrome hiponatremia hipoosmolar hipovolemia koreksi hipokalemia hipoalbuminemia riwayat syok hipovolemik low intake riwayat apendectomy november riwayat operasi kista endometrium	['B-dis', 'O', 'O', 'O', 'O', 'O', 'O', 'B-abb', 'O', 'B-abb', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-dis', 'O', 'B-dis', 'O', 'B-dis', 'B-dis', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'O', 'B-tre', 'B-dis', 'O']
9	periksa fisik bb tb imt kg ku sakit compos mentis td hr rr suhu spo room air mata konjungtiva pucat sklera anikterik leher jvp tingkat kgb raba besar paru vesikular paru rhonki wheezing jantung atur murmur gallop abdomen supel nyeri tekan hepar lien raba bising usus normal ekstremitas akral hangat crt detik edema	['O', 'O', 'B-abb', 'B-abb', 'B-abb', 'B-abb', 'B-abb', 'O', 'B-sym', 'I-sym', 'B-abb', 'B-abb', 'B-abb', 'O', 'B-abb', 'O', 'O', 'B-bod', 'O', 'B-sym', 'B-sym', 'I-sym', 'B-bod', 'B-abb', 'O', 'B-abb', 'O', 'I-tre', 'B-bod', 'B-sym', 'B-bod', 'B-sym', 'B-sym', 'B-bod', 'O', 'B-sym', 'B-sym', 'I-tre', 'O', 'B-sym', 'b-nur', 'I-sym', 'O', 'O', 'O', 'B-bod', 'O', 'B-sym', 'O', 'O', 'B-abb', 'O', 'B-dis']
10	pasien kirim pol hom kemo	['O', 'O', 'B-abb', 'O', 'B-abb']
..	..	…

4. Pemodelan

4.1 Pemodelan NER berbasis Neural Network

Fine-tuning model deep learning untuk NER dan Eksperimen arsitektur hybrid. Pada proses modeling, dataset dibagi menjadi 70% data untuk pelatihan dan 30% data untuk pengujian. Sebelum dilakukan proses pelatihan, perlu membentuk model NER Berbasis Neural Network. Gambar 1 merupakan model NER Berbasis Neural Network yang akan digunakan untuk membentuk model.



Gambar 1. Arsitektur dari model NER berbasis Neural Network

Sebelum membangun model maka pada Gambar 1 dapat dilihat visualisasi arsitektur dari model named entity recognition berbasis IndoBert (4) dan penelitian ini membangun model named entity recognition berbasis IndoBert menggunakan beberapa pendekatan deep learning diantaranya : LSTM, GRU, RNN-biLSTM, RNN-biGRU. Pembangunan model named entity recognition berbasis IndoBert melalui beberapa langkah, diantaranya:

1. Konversi kalimat menjadi urutan, memisahkan tag target dari kalimat, kemudian tokenisasi kalimat menjadi urutan bilangan bulat. Langkah ini dilakukan dengan menggunakan fungsi metode texts_to_sequences. Setiap setence (Tabel C.5) yang akan mengubah setiap kalimat menjadi urutan bilangan bulat, di mana setiap bilangan bulat mewakili kata dalam kalimat.
2. Pemetaan kata ke bilangan bulat, membuat kosakata kata unik beserta indeks bilangan bulatnya. Langkah ini dilakukan dengan menggunakan fungsi fit_on_texts() dari library keras.preprocessing.text yang akan membuat kosakata kata unik beserta indeks bilangan bulatnya. Indeks bilangan bulat ini akan digunakan untuk mewakili kata-kata dalam kalimat.
3. Identifikasi tag NER unik, menemukan tag NER unik yang ada di set pelatihan dan pengujian. Langkah ini dilakukan dengan menggunakan fungsi set() yang akan mengembalikan himpunan semua tag NER unik yang ada di set pelatihan dan pengujian.
4. Konversi target menjadi urutan, proses tokenisasi tag target menjadi urutan bilangan bulat. Langkah ini dilakukan dengan menggunakan fungsi texts_to_sequences() dari library keras.preprocessing.text yang akan mengubah setiap urutan tag menjadi urutan bilangan bulat.
5. Isi baris, memastikan semua urutan memiliki panjang yang sama dengan menambahkan padding. Langkah ini dilakukan dengan menggunakan fungsi pad_sequences() dari library tensorflow.keras.preprocessing.sequence yang akan menambahkan padding ke setiap urutan sehingga semua urutan memiliki panjang yang sama.
6. Jumlah kelas, proses menentukan jumlah tag NER unik (kelas) untuk keluaran model. Langkah ini dilakukan dengan menggunakan fungsi len(), fungsi ini mengembalikan jumlah tag NER unik.
7. Buat model untuk lapisan embedding untuk mewakili kata sebagai vektor padat menggunakan lapisan LSTM bidirectional untuk menangkap konteks di kedua arah. Menggunakan lapisan padat dengan unit keluaran yang sama dengan jumlah tag NER, langkah ini dilakukan dengan menggunakan fungsi Model() dari library keras.models yang membuat model jaringan saraf.
8. Menyiapkan model untuk pelatihan dengan menentukan:
   a. optimizer="adam", algoritma optimasi Adam menggunakan nilai kerugian untuk memperbarui bobot model, mengarahkan model untuk menghasilkan prediksi yang lebih akurat pada iterasi berikutnya.
   b. loss=SparseCategoricalCrossentropy(from_logits=True), fungsi kerugian yang mengukur perbedaan antara prediksi model dan label sebenarnya untuk tugas klasifikasi multi-kelas.
   c. metrics=["accuracy"]:Metrik, mengukur keakuratan prediksi model selama pelatihan.
9. Melatih model dengan data yang disediakan.
   a. train_inputs_final, data masukan untuk pelatihan (kalimat yang sudah diproses)
   b. train_targets_final, label target untuk pelatihan (tak NER yang benar)
   c. epochs=5, Jumlah iterasi pelatihan yang akan dilakukan.

d. validation_data=(test_inputs_final, test_targets_final), data validasi untuk memantau performa model selama pelatihan tanpa mempengaruhi proses pembelajaran model.

Hyperparameters pada Tabel 6, merupakan nilai parameter yang digunakan untuk mengembangan model NER berbasis Neural Network.

Tabel 6. Hyperparameters model berbasis neural network

Parameter	Nilai
text_size	0,3 (70% latih dan 30% uji)
random_state	42
vector size	100
LSTM units	32
epochs	5
optimizer	Adam
loss	SparseCategorialCrossentropy

Hasil pengujian model dengan nilai parameter pada Tabel 6, dapat dilihat pada Tabel 7. Confusion matrix dan classification report pada model NER berbasis Neural Network memiliki nilai bervariasi. Neural network yang digunakan yaitu GRU, bidirectional GRU, LSTM, bidirectional LSTM.

Tabel 7. Evaluasi model neural network

Model	Precision	Recall	F1-Score	Loss	Acuracy
LSTM	0.00494	0.06471	0.00901	0.00354	0.99934
RNN-biLSTM	0.02304	0.06471	0.02483	0.00345	0.99935
GRU	0.00494	0.06471	0.00902	0.00396	0.99937
RNN-biGRU	0.02304	0.06471	0.02483	0.00341	0.99932

Pada Tabel 7, diperlihatkan bahwa hasil evaluasi semua model NER berbasis Neural Network menghasilkan precision sebesar 0.02304 menunjukkan bahwa model hanya benar dalam memprediksi 2.30% label, recall sebesar 0.06471 menunjukkan bahwa model hanya mendeteksi 6.47% label yang benar, f1-score sebesar 0.0248 menunjukkan bahwa model memiliki precision dan recall yang rendah, loss sebesar 0.00345 menunjukkan bahwa model memiliki kesalahan yang rendah, accuracy sebesar 0.99935 menunjukkan bahwa model memiliki akurasi yang sangat tinggi. Secara umum dan dapat model NER berbasis Neural Network memiliki precision, recall, dan f1-score yang rendah, tetapi memiliki accuracy yang sangat tinggi (5). Hal ini menunjukkan bahwa model tersebut sangat baik dalam memprediksi label yang benar, tetapi sering salah mengidentifikasi label yang benar sebagai label yang salah.

4.2 Pemodelan NER berbasis IndoBERT

Penggunaan arsitektur transformers yang merupakan state-of-the-art dari deep learning untuk membangun model named entity recognition menggunakan pretrained dari IndoBERT (6) dan berbasis bidirectional encoder representations from transformers. Peneliti melakukan fine tuning dengan dataset rekam medis elektronik Rumah Sakit Umum Pusat Nasional dr. Cipto Mangunkusumo. Proses dalam membangun model diantaranya:

1. Pemilihan dan eksplorasi data pada dataset ini kolom ‘Sentence’ dan ‘Tag’ dari

dataframe asli. df_final=df[['Sentence','Tag']]

2. Analisis label dari kolom ‘Tag’ untuk memahami jenis tag yang berbeda yang ada. labels = [set([val for sublist in df_final['Tag'].values for val in sublist])]
3. Pengkodean label label2index akan memberikan indeks numerik ke setiap label unik, sehingga lebih mudah bagi model pembelajaran mesin untuk memprosesnya. Sedangkan index2label mempunyai fungsi pemetaan balik untuk mendekripsi indeks kembali ke label.
4. Tokenisasi dan konversi label dengan membagi kalimat menjadi kata-kata individual dan mengkonversi label teks menjadi indeks numerik yang sesuai menggunakan label2index.
5. Pembersihan data dengan memeriksa inkonsistensi, mengidentifikasi baris di mana jumlah token tidak cocok dengan jumlah label, menghapus baris yang tidak konsisten dan memastikan integritas data untuk pelatihan model.
6. Mengonversi dataframe ke kamus dengan mempersiapkan data untuk membuat objek dataset, membuat objek dataset, menggunakan kelas dataset dari perpustakaan datasets untuk membuat dataset terstruktur.
7. Pemisahan dataset menjadi kumpulan pelatihan 70% dan kumpulan pengujian 30% untuk evaluasi model.
8. Tokenisasi dengan memuat tokenizer IndoBERT yang telah dilatih sebelumnya untuk pemrosesan teks bahasa Indonesia.
9. DataCollatorForTokenClassification, berfungsi membuat data collator yang menangani padding dinamis, memastikan panjang input yang konsisten untuk model.
10. Definisi model dengan mempersiapkan model untuk tugas named entity recognition menggunakan model pre-trained IndoBERT.
11. Training arguments dengan mendefinisikan pengaturan untuk proses pelatihan model training_args =TrainingArguments(...).
12. Trainer setup dengan membuat objek Trainer untuk melakukan proses pelatihan dan evaluasi model trainer = Trainer(...).
13. Pada trainer.train() berfungsi memulai proses pelatihan model berdasarkan pengaturan yang telah ditentukan.
14. Pada trainer.save_model("NER_RME") berfungsi menyimpan model yang sudah dilatih ke direktori "NER_RME".
15. Memuat pustaka evaluasi dengan seqeval = evaluate.load ("seqeval") yang dikhususkan untuk mengevaluasi model named entity recognition.
16. Menyiapkan label dengan label_names=[key for key in label2index.keys()] akan nama-nama label dari kamus label2index yang digunakan untuk pengkodean label.
17. Mendefinisikan fungsi evaluasi dengan compute_metrics(p)yang akan menerima prediksi model dan label yang sebenarnya untuk menghitung metrik evaluasi.

Hyperparameters pada Tabel 8, merupakan nilai parameter yang digunakan untuk mengembangan model NER berbasis IndoBERT.

Tabel 8. Hyperparameters model berbasis IndoBERT

Pengujian	Parameter	Nilai
1.	Learning_rate	2e-5
	Train_batch_size	9
	Eval_batch_size	6
	Epochs	5
	Weight_decay	0.01

Pengujian	Parameter	Nilai
2.	Learning_rate	3e-5
	Train_batch_size	9
	Eval_batch_size	6
	Epochs	5
	Weight_decay	0.01
3.	Learning_rate	5e-5
	Train_batch_size	9
	Eval_batch_size	6
	Epochs	5
	Weight_decay	0.01
4.	Learning_rate	5e-5
	Train_batch_size	32
	Eval_batch_size	16
	Epochs	5
	Weight_decay	0.01
5.	Learning_rate	5e-5
	Train_batch_size	6
	Eval_batch_size	3
	Epochs	5
	Weight_decay	0.01

Pada evaluasi model NER berbasis IndoBERT nilai akurasi sebesar 0.99153 atau 99% dalam mengenali entitas bernama, tentu dengan nilai loss sebesar 0.04312 atau 0.4% menunjukkan memiliki kesalahan yang rendah, lihat Tabel 9.

Tabel 9. Tabulasi evaluasi model NER berbasis IndoBERT

No.	Precision	Recall	F1	Loss	Accuracy
1.	0.94982	0.99048	0.96973	0.07344	0.98210
2.	0.96022	0.99524	0.97742	0.06107	0.98634
3.	0.97171	0.99592	0.98366	0.05057	0.99043
4.	0.94748	0.98890	0.96774	0.06859	0.98108
5.	0.97454	0.99705	0.98566	0.04312	0.99153

Pelatihan dan evaluasi model NER berbasis IndoBERT dengan keterbatasan kemampuan perangkat yang baik (lihat Tabel 9), pada tabel tersebut dapat simpulkan dilihat bahwa model 5 memiliki nilai precision sebesar 0.97454 menunjukkan bahwa model hanya salah mengidentifikasi 0.03% label, recall sebesar 0.99705 menunjukkan bahwa model hanya melewatkan 0.01% label yang benar, f1-score sebesar 0.98566 menunjukkan bahwa model memiliki precision dan recall yang sangat tinggi, loss sebesar 0.04312 menunjukkan bahwa model memiliki kesalahan yang rendah, accuracy sebesar 0.99153 menunjukkan bahwa model memiliki akurasi yang sangat tinggi.

Secara umum, model NER berbasis IndoBERT memiliki precision, recall, dan f1-score yang sangat tinggi, serta memiliki loss yang rendah (7). Hal ini menunjukkan bahwa model NER berbasis IndoBERT sangat baik dalam memprediksi label yang benar dan tidak salah mengidentifikasi label yang benar sebagai label yang salah.

Dari perbandingan kedua pendekatan tersebut dalam membangun model NER dapat disimpulkan bahwa pendekatan berbasis IndoBERT secara konsisten mengungguli pendekatan berbasis Neural Network dalam setiap metrik evaluasi.

Pendekatan IndoBERT pada model NER memiliki nilai precision, recall, dan f1-score yang sangat tinggi, menunjukkan kemampuannya dalam mengenali entitas dengan sangat baik, Meskipun model NER berbasis Neural Network (8) memiliki akurasi tinggi, nilai precision dan recall yang rendah menunjukkan bahwa model ini mungkin mengalami kesulitan dalam mengenali entitas dengan akurat, perbedaan yang signifikan antara hasil evaluasi keduanya menunjukkan bahwa pendekatan IndoBERTmemiliki keunggulan dalam tugas named entity recognition dibandingkan dengan Neural Network.

4.3 Penentuan Klasifikasi Penyakit berbasis ICD-10 berbahasa Inggris

Model Penentuan Klasifikasi Penyakit berbasis ICD-10 berbahasa Inggris menggunakan metode Fuzzy Matching, Sentence Transformer dan Keyword Matching. Berikut ini dijelaskan proses pencocokan ICD-10:

4.3.1 Fuzzy Matching

Menggunakan rasio Levenshtein untuk mengukur kedekatan antar string. Setiap entitas dibandingkan dengan deskripsi ICD-10. Skor kemiripan dihitung, lalu dipilih ICD-10 dengan skor tertinggi.Contoh: teks pasien “apendisitis” dan label ICD “appendicitis” memiliki jarak edit kecil sehingga skor fuzzy tinggi. Contoh lain misalnya "nyeri perut kanan" bisa dibandingkan dengan ICD-10 "Right lower quadrant pain".

Skor: Hasilnya berupa angka 0–100, semakin tinggi semakin mirip. Threshold: default-nya threshold=70. Artinya, hanya hasil dengan kemiripan ≥70 yang dianggap relevan. Metode ini efektif untuk kasus transliterasi, ejaan lokal, atau variasi singkatan.

4.3.2 Semantic Matching

Semantic matching adalah proses mencocokkan informasi berdasarkan kesamaan makna dan niat di balik teks, bukan hanya berdasarkan kata kunci yang sama (9). Ini melibatkan pemahaman konteks dan arti sebenarnya dari sebuah input untuk menemukan kecocokan yang lebih akurat dan relevan, yang merupakan dasar dari penelusuran semantik dan berbagai aplikasi kecerdasan buatan lainnya. Jadi, dalam proses pencocokan semantik, kami melatih jaringan yang menghasilkan skor kesamaan yang tinggi ketika pertanyaannya serupa, dan skor kesamaan yang rendah ketika pertanyaannya berbeda.

Berikut ini cara kerja semantic matching:

- Memahami Makna: Sistem semantic matching tidak hanya melihat kata-kata, tetapi juga menganalisis hubungan antara kata, frasa, dan kalimat untuk memahami makna keseluruhan.
- Mempertimbangkan Konteks: Konteks sangat penting. Kata yang sama bisa memiliki arti yang berbeda tergantung pada kalimatnya. Semantic matching berusaha memahami konteks ini.
- Analisis Niat Pengguna: Dalam penelusuran, sistem akan mencoba memahami niat di balik kueri pengguna.

4.3.3 Keyword Matching

Proses ini memeriksa apakah token kunci dalam teks pasien muncul dalam deskripsi ICD-10 (10). Jika ditemukan, maka kandidat ICD tersebut diberi skor lebih tinggi. Contohnya, kata kunci “gastritis” pada teks pasien langsung cocok dengan ICD yang mengandung kata “gastritis”. Contoh lain Misalnya, jika entitas mengandung kata "perut" maka dicari ICD-10 yang mengandung kata "abdomen" atau "stomach". Metode ini bekerja cepat tetapi sensitif terhadap sinonim dan variasi bahasa.

Tahapan proses pencocokan dengan keyword matching sebagai berikut:

1. Daftar kata dari deskripsi ICD-10 dibagi menjadi token (kata).
2. Setiap kata yang panjangnya lebih dari 3 huruf dimasukkan ke kamus keyword_to_icd.
3. Entitas pasien juga dipecah menjadi kata, lalu dicek apakah ada kata yang cocok dengan kamus.
4. Jika ada beberapa kata yang cocok, sistem menghitung frekuensi kode ICD-10 yang muncul, dan memilih yang paling sering muncul.

Skor: Jumlah kata kunci yang cocok. Threshold: Tidak eksplisit dalam bentuk angka (seperti fuzzy/semantic), tetapi secara implisit threshold-nya adalah ≥1 kata kunci yang cocok. Jika tidak ada kata sama sekali yang cocok, hasilnya tidak diterima.

4.3.4 Metrik Evaluasi

Fungsi evaluasi yang membandingkan hasil tiga metode matching, menampilkan skor, lalu membuat visualisasi perbandingan.

```
=== Evaluation Results untuk 100 Patient Pertama ===
                match_rate  avg_score  total_matches
fuzzy_match           47.0  90.638298           47.0
keyword_match         38.0   5.447368           38.0
semantic_match        80.0  81.868625           80.0
```

Masing-masing metrik memberikan perspektif yang berbeda tentang efektivitas dan kualitas hasil pencocokan.

Match_rate: Metrik ini merepresentasikan persentase pasien yang berhasil mendapatkan setidaknya satu pasangan ICD-10 menggunakan metode tertentu, baik itu Fuzzy Matching, Keyword Matching, maupun Semantic Matching. Perhitungannya dilakukan dengan membagi jumlah pasien yang memiliki minimal satu hasil cocok dengan jumlah keseluruhan pasien, kemudian dikalikan 100. Semakin tinggi match rate, semakin banyak pasien yang berhasil diberikan label ICD-10 oleh metode tersebut. Dengan kata lain, match rate menunjukkan seberapa sering suatu metode dapat menemukan *relevan mapping*.

Average Score (avg_score): Metrik ini mencerminkan kualitas rata-rata dari hasil pencocokan yang berhasil. Cara perhitungan skor bergantung pada metode yang digunakan:

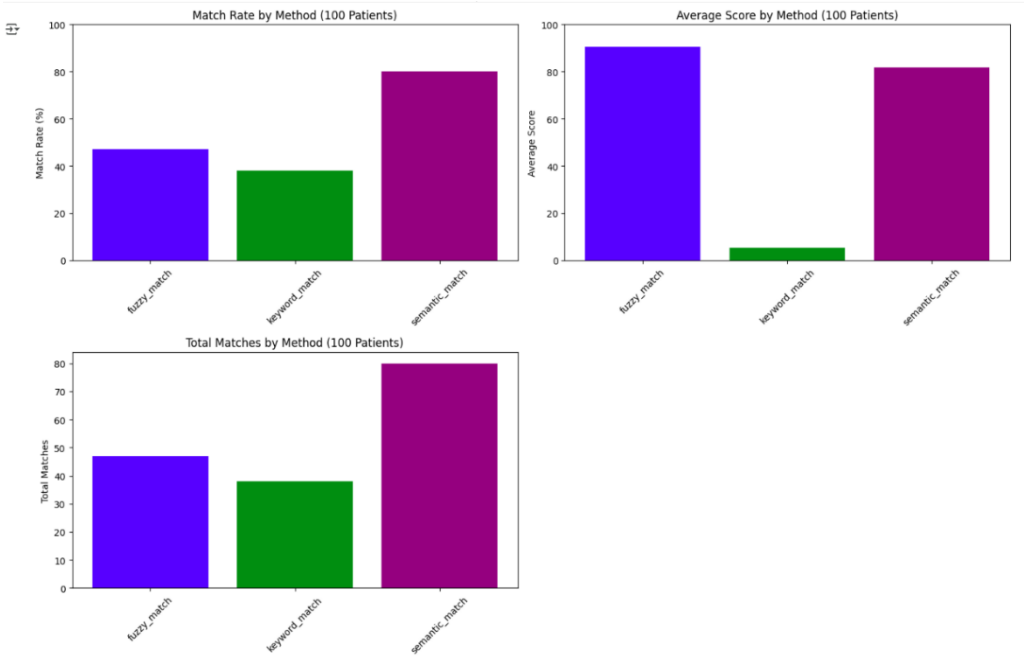
- Pada Fuzzy Matching, skor dihitung menggunakan *fuzzy similarity* berbasis Levenshtein distance, dengan rentang 0 hingga 100. Nilai tinggi berarti teks pasien sangat mirip secara string dengan deskripsi ICD-10.
- Pada Keyword Matching, skor dihitung berdasarkan jumlah kata kunci yang cocok. Namun, dalam implementasi saat ini, jika ada satu kata kunci yang cocok maka skor diberikan 1.0, sedangkan jika tidak ada maka nilainya 0. Sehingga nilai rata-rata dari metode ini kurang informatif karena tidak membedakan antara kecocokan parsial dan kecocokan penuh.

- Pada Semantic Matching, skor berasal dari *cosine similarity* antar embedding kalimat yang dihitung oleh model Sentence Transformer. Rentangnya 0–1, lalu diskalakan menjadi 0–100 untuk memudahkan interpretasi. Skor tinggi menunjukkan bahwa secara semantik kalimat pasien dan deskripsi ICD-10 memiliki makna yang sangat dekat.

Total Matches: Metrik ini hanya menghitung jumlah total pencocokan ICD-10 yang berhasil ditemukan suatu metode pada seluruh pasien. Angka ini menjadi dasar perhitungan *match rate*. Misalnya, jika dari 10 pasien terdapat 8 yang berhasil diberikan ICD-10 oleh metode fuzzy, maka *total matches* = 8 dan *match rate* = 80%.

Secara ringkas, *match rate* menunjukkan frekuensi keberhasilan metode menemukan pasangan ICD-10, sedangkan *avg_score* menunjukkan kualitas rata-rata hasil pencocokan tersebut. *Total matches* memberikan jumlah absolut pasangan yang berhasil.

Visualization

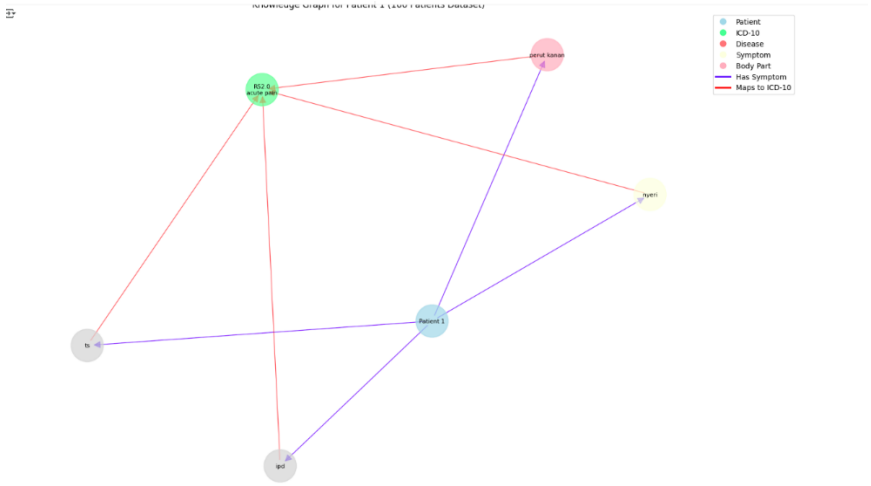


Summary Report

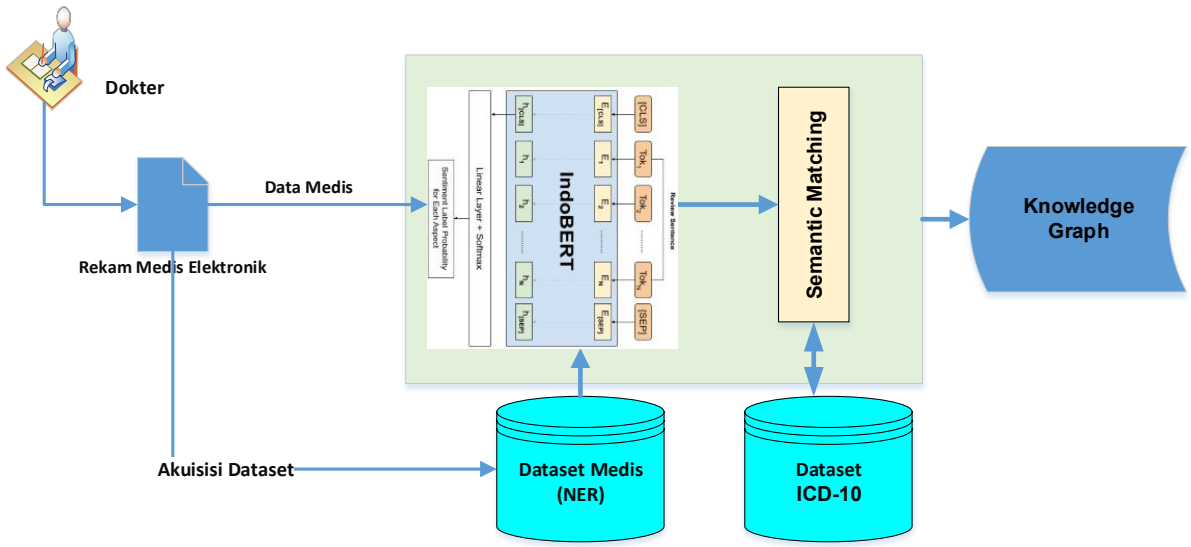
```
=== SUMMARY REPORT (100 Patient Pertama) ===
Total patients processed: 100
Best matching method based on average score: fuzzy_match
Highest match rate: semantic_match

=== SAMPLE MAPPED RESULTS (10 Patient Pertama) ===
   patient_id fuzzy_match_code keyword_match_code semantic_match_code
0           1              None               None               R52.0
1           2              None                A90                A90
2           3               K35               None                K37
3           4             R01.1              R01.1               B81.0
4           5              None               None               N48.3
5           6              None               None               N48.3
6           7              None               None                None
7           8             I86.4              I86.4              I86.4
8           9             R06.2              R06.2                A46
9          10              None               None                None
```

Knowledge Graph

Untuk tiga pasien pertama, hasil matching divisualisasikan dalam bentuk knowledge graph (11). Dengan NetworkX, node pasien, gejala, dan ICD-10 dibuat dengan warna berbeda, lalu matplotlib digunakan untuk menggambar graf. Misalnya, node "Patient 1" terhubung ke node "nyeri perut kanan", yang kemudian terhubung ke node ICD-10 "R10.31 Right lower quadrant pain".



Hasil eksperimen dengan berbagai pendekatan, teknologi, metode dan parameter, maka Model akhir yang dihasilkan adalah Model NER berbasis IndoBERT dan Semantic Matching. Model penentuan penyakit berbasis ICD-10 dapat dilihat pada Gambar 2.



Gamabr 2. Penentuan Klasifikasi Penyakit berbasis ICD-10 berbahasa Inggris

.

D. **STATUS LUARAN**: Tuliskan jenis, identitas dan status ketercapaian setiap luaran wajib dan luaran tambahan (jika ada) yang dijanjikan. Jenis luaran dapat berupa publikasi, perolehan kekayaan intelektual, atau luaran lainnya yang telah dijanjikan pada proposal. Uraian status luaran harus didukung dengan bukti kemajuan ketercapaian luaran sesuai dengan luaran yang dijanjikan. Lengkapi isian jenis luaran yang dijanjikan serta mengunggah bukti dokumen ketercapaian luaran melalui BIMA.

Pada saat ini target luaran masih dalam proses pembuatan, yaitu proses penyusunan paper publikasi. Target publikasi masih sesuai dengan target awal pada saat proposal. Paper publikasi dapat dilihat pada file lampiran yang terpisah.

E. PERAN MITRA: Tuliskan realisasi kerjasama dan kontribusi Mitra baik *in-kind* maupun *in-cash* serta mengunggah bukti dokumen pendukung sesuai dengan kondisi yang sebenarnya. Bukti dokumen realisasi kerjasama dengan Mitra dapat diunggah melalui BIMA.

Catatan:
Bagian ini wajib diisi untuk penelitian terapan, untuk penelitian dasar (KATALIS, Fundamental, Pascasarjana, dan Dosen Pemula) boleh mengisi bagian ini (tidak wajib) jika melibatkan mitra dalam pelaksanaan penelitiannya

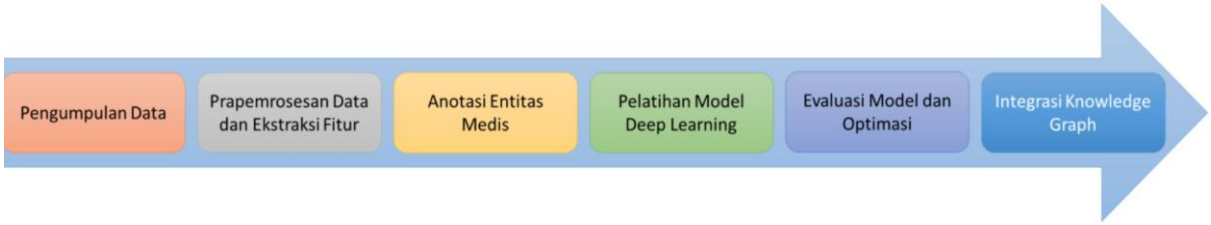
Rumah sakit yang dijadikan mitra dalam penelitian adalah Rumah Sakit Umum Pusat Nasional dr. Cipto Mangunkusumo. Rumah sakit Mitra di sini digunakan untuk mendapatkan data awal dan mendapatkan dokter umum dan dokter spesialis. Peran dokter umum dan dokter spesialis adalah untuk melakukan validasi data dan pengujian model yang dihasilkan.

F. KENDALA PELAKSANAAN PENELITIAN: Tuliskan kesulitan atau hambatan yang dihadapi selama melakukan penelitian dan mencapai luaran yang dijanjikan, termasuk penjelasan jika pelaksanaan penelitian dan luaran penelitian tidak sesuai dengan yang direncanakan atau dijanjikan.

Kendala yang dihadapi saat pelaksanaan penelitian adalah data ICD-10 dalam bentuk bahasa inggris. Hal tersebut menyebabkan adanya pekerjaan tambahan, yaitu menambahkan metode untuk menterjemahkan ICD-10 dalam bahasa Indonesia. Namun, kendala tersebut tidak sampai mempengaruhi capaian dari penelitian yang dilakukan

G. RENCANA TAHAPAN SELANJUTNYA: Tuliskan dan uraikan rencana penelitian selanjutnya berdasarkan indikator luaran yang telah dicapai, rencana realisasi luaran wajib yang dijanjikan dan tambahan (jika ada) di tahun berikutnya serta *roadmap* penelitian keseluruhan. Pada bagian ini diperbolehkan untuk melengkapi penjelasan dari setiap tahapan dalam metoda yang akan direncanakan termasuk jadwal berkaitan dengan strategi untuk mencapai luaran seperti yang telah dijanjikan dalam proposal. Jika diperlukan, penjelasan dapat juga dilengkapi dengan gambar, tabel, diagram, serta pustaka yang relevan. Jika laporan kemajuan merupakan laporan pelaksanaan tahun terakhir, pada bagian ini dapat dituliskan rencana penyelesaian target yang belum tercapai.

Berdasarkan Gambar 3, penelitian terdapat 6 kegiatan atau tahapan, dari 6 kegiatan yang ada kegiatan keenam (terakhiri) yang sedang proses penyelesaian. Kegiatan tersebut adalah Integrasi Knowledge Graph. Dalam kegiatan tersebut meliputi proses validasi dan integrasi. Integrasi knowledge graph merupakan proses kemudian pengujian knowledge graph yang dihasilkan dan menghubungkan entitas dengan UMLS/SNOMED-CT untuk memperkaya konteks, kemudian melakukan pengujian. Kemudian kegiatan terakhir adalah melakukan Validasi Klinis, yaitu melakukan pengujian model yang dihasilkan di lingkungan rumah sakit.



Gambar 3. Diagram Tahapan Penelitian

.

H. DAFTAR PUSTAKA: Penyusunan Daftar Pustaka berdasarkan sistem nomor sesuai dengan urutan pengutipan. Hanya pustaka yang disitasi pada laporan kemajuan yang dicantumkan dalam Daftar Pustaka.

.1. Turki H, Etori NA, Hadj Taieb MA, Omotayo AH, Emezue CC, Ben Aouicha M, et al. Text Categorization Can Enhance Domain-Agnostic Stopword Extraction. In 2025. p. 122–32. Available from: https://link.springer.com/10.1007/978-3-031-85067-7_12

2. Khan MA, Anwar S, Abbas M, Aneeq M, de Jong F, Ayaz M, et al. Impacts of climate change on cotton production and advancements in genomic approaches for stress resilience enhancement. J Cott Res [Internet]. 2025 May 12;8(1):17. Available from: https://jcottonres.biomedcentral.com/articles/10.1186/s42397-025-00223-3

3. Bosma JS, Dercksen K, Builtjes L, André R, Roest C, Fransen SJ, et al. The DRAGON benchmark for clinical NLP. npj Digit Med [Internet]. 2025 May 17;8(1):289. Available from: https://www.nature.com/articles/s41746-025-01626-x

4. Bryan Wilie, Karissa Vincentio GIW. IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding. In: Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing. 2020. p. 843–57.

5. Chen J, Xi X, Sheng VS, Cui Z. Randomly Wired Graph Neural Network for Chinese NER. Expert Syst Appl [Internet]. 2023 Oct;227:120245. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0957417423007479

6. Nabiilah GZ, Prasetyo SY, Izdihar ZN, Girsang AS. BERT base model for toxic comment analysis on Indonesian social media. Procedia Comput Sci [Internet]. 2023;216:714–21. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1877050922022669

7. Ramos-Flores O, Gómez-Adorno H, Vazquez M, De Ita R, Mimiaga-Morales JM, Campos-Campechano FDA, et al. Manual annotation of Robson criteria and obstetric entities: Inter-annotator agreement and initial NER models implementation. Comput Biol Med [Internet]. 2025 Oct;197:110964. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0010482525013162

8. Mengliev D, Barakhnin V, Eshkulov M, Ibragimov B, Madirimov S. A comprehensive dataset and neural network approach for named entity recognition in the Uzbek language. Data Br [Internet]. 2025 Feb;58:111249. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2352340924012113

9. Li M, Gao Y, Zhao H, Li R, Chen J. Progressive semantic aggregation and structured cognitive enhancement for image–text matching. Expert Syst Appl [Internet]. 2025 May;274:126943. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0957417425005652

10. Swain OP, Hemanth H, Saran P, Kothandaraman M, Ravi L, Sailor H, et al. Robust and

efficient keyword spotting using a bidirectional attention LSTM. Int J Speech Technol [Internet]. 2023 Dec 11;26(4):919–31. Available from: https://link.springer.com/10.1007/s10772-023-10067-4

11. Zhuang J, Yang J, Li W, Chen J, Zheng Y, Chen Z. Large model for fault diagnosis of industrial equipment based on a knowledge graph construction. Appl Soft Comput [Internet]. 2025 Dec;185:113936. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1568494625012499