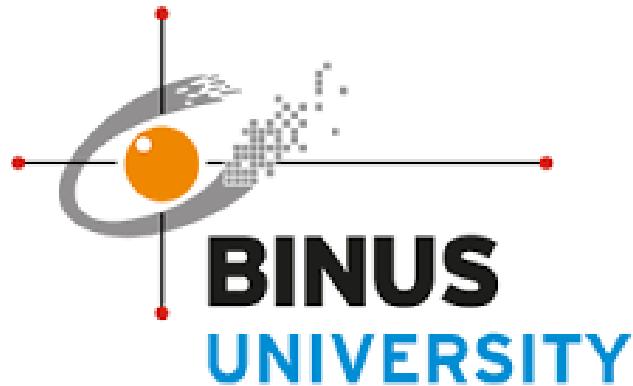


## LAPORAN PROJECT AOL DEEP LEARNING

# Sentiment-Aware Named Entity Recognition for Analyzing Food and Service Quality in Restaurants



Kelas : LB01

Dosen Pembimbing:

Wawan Cenggoro, S.Kom., M.TI D5544

Disusun oleh:

Rainer Yesaya Villareal | 2702275730

William Hartanto | 2702274532

FAKULTAS ILMU KOMPUTER – COMPUTER SCIENCE  
UNIVERSITAS BINA NUSANTARA  
JAKARTA 2025

## ABSTRACT

Volume ulasan daring yang masif menghadirkan tantangan dalam analisis manual reputasi restoran. Laporan ini menguraikan pengembangan sistem *Sentiment-Aware Named Entity Recognition* (SA-NER) berbasis *Deep Learning* untuk mengotomatisasi analisis aspek makanan dan layanan. Menggunakan arsitektur DistilBERT yang di-*fine-tune*, sistem ini mengintegrasikan *Named Entity Recognition* (NER) untuk identifikasi entitas dan Analisis Sentimen untuk penentuan polaritas.

Melalui pengujian menggunakan *MIT Restaurant Corpus* dan dataset sentimen terkait, model mencapai **F1-Score NER sebesar 80.62%** dan **akurasi sentimen 93.0%**. Pendekatan ini terbukti efektif menangani kalimat majemuk dengan sentimen campuran. Hasil akhir diimplementasikan dalam aplikasi web berbasis Streamlit, yang mendemonstrasikan kemampuan sistem dalam membantu pemilik restoran memetakan kekuatan dan kelemahan operasional secara efektif.

**Kata kunci:** *Sentiment-Aware NER, DistilBERT, Deep Learning, Analisis Sentimen, Ulasan Restoran, Natural Language Processing.*

## BAB I

### 1. PENDAHULUAN

#### 1.1. Latar Belakang

Transformasi digital dalam industri kuliner telah menjadikan ulasan daring (*online reviews*) sebagai aset strategis. Platform seperti Google Maps dan Zomato kini menampung miliaran ulasan pengguna yang berfungsi sebagai "word-of-mouth" digital. Bagi pemilik restoran, data ini menawarkan wawasan krusial mengenai operasional dan layanan. Namun, volume data yang masif dan sifat teks yang tidak terstruktur menjadikan analisis manual menjadi tidak praktis, mahal, dan rentan terhadap bias subjektif.

Tantangan utama dalam pemrosesan ulasan otomatis adalah kompleksitas kalimat yang seringkali memuat sentimen campuran. Sebagai contoh, kalimat "*Steaknya luar biasa, tetapi pelayanannya lambat*" mengandung polaritas positif dan negatif sekaligus pada entitas yang berbeda. Metode Analisis Sentimen tradisional yang mengklasifikasikan polaritas pada tingkat kalimat atau dokumen gagal menangkap granularitas ini, sehingga sering kali menghilangkan informasi kritis mengenai aspek spesifik yang perlu diperbaiki.

Untuk mengatasi kesenjangan tersebut, diperlukan pendekatan *Sentiment-Aware Named Entity Recognition* (SA-NER) yang mampu mendeteksi entitas spesifik sekaligus sentimen yang melekat padanya. Penelitian ini memanfaatkan kemajuan *Deep Learning* menggunakan arsitektur DistilBERT. Model ini dipilih karena menawarkan kompromi optimal antara performa prediktif arsitektur Transformer (seperti BERT) dan efisiensi komputasi, yang krusial untuk implementasi

bisnis. Proyek ini bertujuan membangun sistem *end-to-end* yang mengolah ulasan mentah menjadi wawasan visual yang dapat ditindaklanjuti.

### **1.2. Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan, proyek ini mengidentifikasi beberapa masalah inti yang perlu diselesaikan:

- a. Bagaimana sistem dapat secara otomatis membedakan antara teks deskriptif biasa dan entitas penting dalam ulasan yang penuh dengan noise linguistik?
- b. Bagaimana sistem dapat secara akurat memetakan sentimen ke entitas yang tepat dalam kalimat majemuk yang mengandung opini campuran (mixed opinions)?
- c. Apakah model yang lebih ringan seperti DistilBERT mampu mempertahankan tingkat akurasi yang memadai untuk tugas ganda dibandingkan dengan model yang lebih besar?

### **1.3. Tujuan**

Tujuan utama dari penelitian dan pengembangan sistem ini adalah:

- a. Membangun model yang mampu mengenali aspek spesifik restoran (seperti makanan atau layanan) sekaligus menentukan sentimen (positif/negatif) dari ulasan pelanggan secara otomatis.
- b. Menciptakan sistem yang dapat mengubah data ulasan mentah menjadi informasi yang terstruktur dan rinci.
- c. Menguji tingkat akurasi dan keandalan sistem untuk memastikan hasil analisis yang dihasilkan valid dan tepat sasaran.
- d. Menyediakan antarmuka visual (aplikasi web) yang mudah digunakan, sehingga dapat memahami hasil analisis dengan cepat untuk mengetahui kekuatan dan kelemahan restoran guna meningkatkan kualitas layanan.

## **BAB II**

### **2. KAJIAN TEORI**

#### **2.1. Deep Learning dalam Pemrosesan Bahasa Alami (NLP)**

*Natural Language Processing* (NLP) adalah cabang dari kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa manusia. Secara historis, NLP bergantung pada sistem berbasis aturan (*rule-based systems*) dan model statistik sederhana seperti *Bag-of-Words* atau *N-Grams*. Namun, pendekatan ini memiliki keterbatasan fatal dalam menangkap konteks, makna semantik, dan urutan kata yang panjang.

*Deep Learning* telah merevolusi NLP dengan memperkenalkan representasi vektor terdistribusi (*word embeddings*) dan arsitektur jaringan saraf yang dalam. Model seperti *Recurrent Neural Networks* (RNN) dan *Long Short-Term Memory* (LSTM) sempat menjadi standar karena kemampuannya memproses data sekuesial. Namun, mereka masih menderita masalah komputasi yang lambat (karena pemrosesan berurutan) dan kesulitan dalam menangkap dependensi jarak jauh dalam kalimat yang sangat panjang.

## 2.2. Arsitektur Transformer dan Mekanisme Attention

Diperkenalkan oleh Vaswani et al. pada tahun 2017, arsitektur Transformer mengubah lanskap NLP secara fundamental. Transformer meninggalkan mekanisme rekurensi (RNN) sepenuhnya dan mengandalkan mekanisme yang disebut Self-Attention. Mekanisme ini memungkinkan model untuk menimbang pentingnya setiap kata dalam kalimat terhadap kata lainnya secara simultan, tanpa memandang jarak posisinya.<sup>3</sup>

## 2.3. BERT (Bidirectional Encoder Representations from Transformers)

BERT, yang dikembangkan oleh Google, adalah implementasi *encoder stack* dari Transformer. Inovasi utama BERT adalah pelatihannya yang bersifat bidirectional (dua arah). Model bahasa sebelumnya biasanya membaca teks dari kiri-ke-kanan atau kanan-ke-kiri. BERT membaca seluruh urutan input sekaligus.

BERT dilatih dengan dua tujuan utama (*pre-training objectives*):

1. **Masked Language Modeling (MLM):** Menyembunyikan sebagian kata dalam kalimat (misalnya 15%) dan memaksa model untuk memprediksi kata tersebut berdasarkan konteks sekitarnya. Ini mengajarkan model pemahaman kontekstual yang mendalam.
2. **Next Sentence Prediction (NSP):** Memprediksi apakah kalimat B secara logis mengikuti kalimat A. Ini mengajarkan model hubungan antar-kalimat.<sup>4</sup>

## 2.4. DistilBERT: Efisiensi Melalui Knowledge Distillation

Meskipun BERT memiliki performa *state-of-the-art*, ukurannya yang besar (110 juta parameter untuk BERT-Base) membuatnya lambat dan mahal untuk *deployment* di lingkungan produksi nyata. DistilBERT (*Distilled BERT*) hadir sebagai solusi yang memanfaatkan teknik Knowledge Distillation untuk memangkas ukuran model sebesar 40% dan mempercepat inferensi hingga 60%, sambil mempertahankan 97% kinerja BERT pada benchmark GLUE.<sup>2</sup>

### 2.4.1. Konsep Knowledge Distillation

Knowledge Distillation adalah teknik kompresi model di mana model kecil ("Student" - DistilBERT) dilatih untuk meniru perilaku model besar yang sudah terlatih ("Teacher" - BERT). Alih-alih hanya belajar dari label keras (hard targets atau gold labels) dari dataset pelatihan, Student belajar dari probabilitas lunak (soft targets) yang dihasilkan oleh Teacher.

Dalam keluaran softmax Teacher, probabilitas kelas yang salah seringkali mengandung informasi semantik yang kaya. Misalnya, jika model memprediksi gambar "kucing", probabilitas kecil untuk menjadi "anjing" mungkin jauh lebih tinggi daripada menjadi "mobil". Informasi relasional antar-kelas ini disebut *dark knowledge*. DistilBERT menangkap pengetahuan ini dengan menggunakan parameter suhu pada fungsi softmax untuk memperhalus distribusi probabilitas, sehingga Student dapat mempelajari nuansa generalisasi yang dimiliki Teacher.<sup>2</sup>

### 2.4.2. Arsitektur Model: Student vs Teacher

Secara struktural, DistilBERT adalah hasil kompresi arsitektur BERT dengan modifikasi berikut:

- Pengurangan Lapisan (Layer Pruning): Jumlah lapisan Transformer dikurangi dari 12 (pada BERT-Base) menjadi 6.
- Inisialisasi Cerdas: Bobot lapisan DistilBERT diinisialisasi dengan mengambil satu dari setiap dua lapisan BERT Teacher. Ini memastikan struktur fitur yang sudah dipelajari Teacher ditransfer langsung ke Student.
- Penyederhanaan Komponen: *Token-type embeddings* dan lapisan *pooler* dihapus untuk efisiensi lebih lanjut.
- Retensi Dimensi: Dimensi tersembunyi (*hidden size*) tetap dipertahankan sebesar 768, memastikan kapasitas representasi model tetap kuat.<sup>5</sup>

#### 2.4.3. Fungsi Kerugian Tiga Komponen (Triple Loss Function)

Proses pelatihan DistilBERT menggunakan fungsi kerugian gabungan yang mengoptimalkan tiga tujuan sekaligus:

1. Distillation Loss: Menggunakan KL Divergence untuk mengukur perbedaan antara distribusi probabilitas lunak Teacher dan Student. Ini adalah inti dari transfer pengetahuan.
2. Masked Language Modeling Loss: Fungsi kerugian standar *cross-entropy* terhadap label asli (teks yang di-mask). Ini menjaga kemampuan model memahami tata bahasa dan konteks dasar.
3. Cosine Embedding Loss: Memastikan arah vektor representasi tersembunyi (*hidden states*) yang dihasilkan Student selaras (kolinear) dengan vektor Teacher. Ini meningkatkan konsistensi representasi fitur antar kedua model.<sup>2</sup>

Kombinasi efisiensi komputasi dan retensi akurasi ini menjadikan DistilBERT tulang punggung yang ideal untuk sistem SA-NER ini, memungkinkan analisis ulasan yang cepat dan responsif tanpa memerlukan infrastruktur GPU kelas atas.<sup>6</sup>

#### 2.5. Named Entity Recognition (NER) dan Skema Tagging BIO

*Named Entity Recognition* (NER) adalah sub-tugas ekstraksi informasi yang bertujuan mengklasifikasikan entitas bernama dalam teks ke dalam kategori yang telah ditentukan (seperti Makanan, Fasilitas, Layanan). Karena teks adalah data sekuensial, NER diperlakukan sebagai masalah *sequence labeling*.

Untuk melatih model NER, setiap kata (token) dalam kalimat harus diberi label. Skema yang paling umum digunakan dan diterapkan dalam proyek ini adalah skema BIO (*Begin*, *Inside*, *Outside*)<sup>7</sup>:

- B-: Menandai awal dari sebuah entitas (misalnya, B-Dish untuk "Nasi").
- I-: Menandai kelanjutan dari entitas yang sama (misalnya, I-Dish untuk "Goreng" dalam frasa "Nasi Goreng").
- O: Menandai token yang bukan merupakan entitas (*Outside*).

Penggunaan skema ini memungkinkan model untuk membedakan batas antara dua entitas sejenis yang muncul berurutan (misalnya, "Saya memesan Burger Pizza"). Tanpa tag 'B', model mungkin menganggap "Burger Pizza" sebagai satu entitas hidangan.

## 2.6. Analisis Sentimen (Sentiment Analysis)

Analisis sentimen adalah proses mengidentifikasi orientasi emosional dalam teks. Dalam konteks ulasan produk atau layanan, ini biasanya dikategorikan sebagai Positif, Negatif, atau Netral. Proyek ini menggunakan pendekatan *supervised learning* dengan *fine-tuning* DistilBERT untuk klasifikasi urutan (*sequence classification*).

Berbeda dengan metode leksikon yang hanya menghitung bobot kata positif/negatif, model berbasis Transformer memanfaatkan token (Classification Token) yang disisipkan di awal kalimat. Representasi vektor dari token setelah melewati seluruh lapisan Transformer dianggap mengandung representasi semantik agregat dari seluruh kalimat, yang kemudian diumpulkan ke lapisan klasifikasi (*fully connected layer*) untuk menentukan probabilitas kelas sentimen.<sup>9</sup>

# BAB III

## 3. DATASET

Keberhasilan model *Deep Learning* sangat bergantung pada kualitas dan relevansi data yang digunakan. Proyek ini memanfaatkan dua dataset terpisah untuk melatih komponen NER dan Analisis Sentimen secara independen.

### 3.1. Dataset NER:

MIT Restaurant Corpus Untuk ekstraksi entitas, digunakan MIT Restaurant Corpus yang merupakan standar dalam domain kuliner. Dataset ini berisi kalimat query pencarian restoran dengan format BIO. Label entitas yang dicakup sangat relevan untuk analisis kualitas, meliputi: Rating, Amenity, Cuisine, Dish, Location, Price, dan Hours. Dataset ini dipilih karena menyediakan variasi linguistik dan terminologi spesifik yang krusial untuk pengenalan konteks restoran.

### 3.2. Dataset Analisis Sentimen:

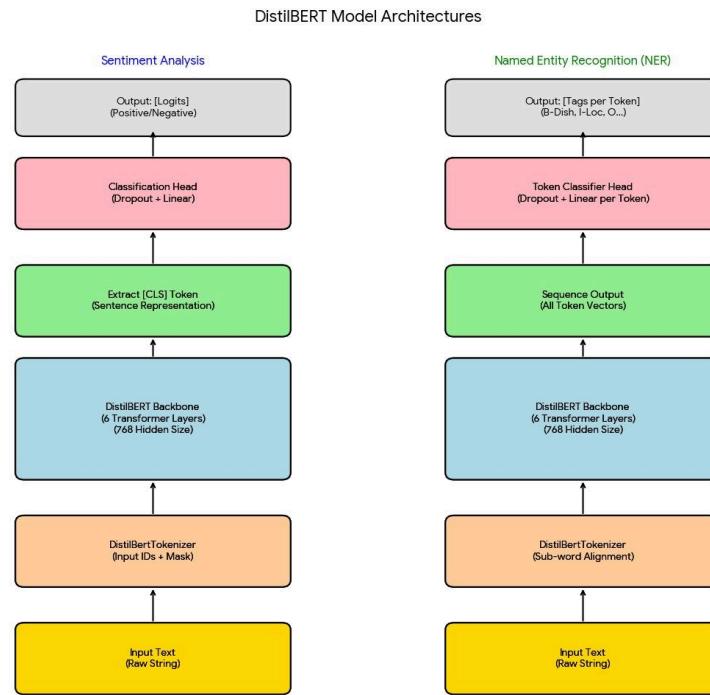
Restaurant Reviews Untuk klasifikasi polaritas, digunakan dataset Restaurant-Reviews-Sentiment-Analysis yang bersumber dari repositori publik. Dataset ini berisi ulasan opini naratif dengan label biner: 1 (Positif) dan 0 (Negatif). Berbeda dengan MIT Corpus, karakteristik linguistik dataset ini bersifat informal dan ekspresi emosional yang merepresentasikan data ulasan dunia nyata. Volume data dinilai memadai untuk melakukan fine-tuning yang efektif pada model DistilBERT.

## BAB IV

### 4. MODELING & METODOLOGI

Bab ini merinci arsitektur sistem, strategi *preprocessing*, konfigurasi pelatihan, dan protokol evaluasi yang diterapkan untuk membangun sistem SA-NER.

#### 4.1. Desain Arsitektur Sistem



Kami mengadopsi pendekatan Pipeline Terpisah (Decoupled Pipeline). Alih-alih melatih satu model *multi-task* yang kompleks, kami melatih dua instansi model DistilBERT secara terpisah: satu didedikasikan untuk NER dan satu untuk Klasifikasi Sentimen. Keputusan ini didasarkan pada pertimbangan bahwa representasi fitur yang optimal untuk mendeteksi batas entitas (sintaksis) mungkin berbeda dengan representasi untuk mendeteksi emosi (semantik).

Arsitektur sistem saat inferensi bekerja secara sekuensial:

1. Input: Teks ulasan mentah.
2. Preprocessing & Segmentasi: Teks dibersihkan dan dipecah menjadi segmen-segmen logis (klausa/kalimat).
3. Model 1 (NER): Memindai setiap segmen untuk mengekstrak entitas (Makanan, Layanan).
4. Model 2 (Sentiment): Menganalisis polaritas sentimen dari segmen tempat entitas ditemukan.
5. Agregasi: Menggabungkan hasil (Entitas + Sentimen) menjadi wawasan terstruktur.

#### 4.2. Preprocessing Data

Merujuk pada detail teknis di Appendix A dokumen referensi <sup>1</sup>, *preprocessing* dilakukan secara spesifik untuk masing-masing model:

#### 4.2.1. Preprocessing untuk NER

Integritas posisi kata adalah prioritas utama.

- **Pembersihan Minimal:** Menghapus karakter *non-printable* atau *encoding* yang rusak, tetapi mempertahankan tanda baca dan kapitalisasi (jika model *cased*) karena sering menjadi indikator nama entitas (Proper Noun).
- **Tokenisasi & Alignment:** Kalimat dipecah menjadi token menggunakan DistilBertTokenizerFast. Fungsi khusus dibuat untuk menyelaraskan daftar label BIO asli dengan daftar token hasil WordPiece. Token spesial dan ditambahkan, dan label `_PAD_` digunakan untuk menyamakan panjang urutan dalam satu *batch*.

#### 4.2.2. Preprocessing untuk Analisis Sentimen

Fokus utama adalah pada "sinyal" emosi.

- **Pembersihan Teks (Cleaning):** Mengubah teks menjadi huruf kecil (*lowercase*), menghapus karakter non-alfabet (angka, simbol khusus), dan menghapus *stopwords* (kata umum seperti "the", "and", "is") yang tidak membawa muatan sentimen signifikan. Ini membantu model fokus pada kata sifat (*adjectives*) dan kata kerja yang deskriptif.<sup>1</sup>
- **Tokenisasi:** Mengubah kalimat menjadi urutan ID token dan *Attention Mask*. *Attention Mask* (bernilai 1 untuk token asli, 0 untuk padding) sangat penting agar mekanisme *Self-Attention* model tidak memproses area kosong.<sup>15</sup>

### 4.3. Konfigurasi Model (Modelling)

Kami menggunakan pustaka Hugging Face Transformers untuk memuat model *pre-trained* distilbert-base-uncased.

#### 4.3.1. Model NER (Token Classification)

- Base Model: DistilBERT (6 layer, 768 hidden dimension).
- Classification Head: Sebuah lapisan *Dropout* (untuk regularisasi) diikuti oleh lapisan *Linear* dengan dimensi output (jumlah label BIO unik dalam dataset).
- Output: Distribusi probabilitas untuk setiap kelas label BIO pada setiap token.

#### 4.3.2. Model Sentimen (Sequence Classification)

- Base Model: DistilBERT.
- Pooling Strategy : Menggunakan representasi vektor sebagai ringkasan semantik kalimat.
- Classification Head: Lapisan *Linear* dengan dimensi output 2 (Positif, Negatif).
- Fungsi Aktivasi: Logits mentah dimasukkan ke fungsi *Softmax* atau *Sigmoid* untuk mendapatkan probabilitas.<sup>9</sup>

### 4.4. Strategi Pelatihan (Training)

Dataset dibagi dengan rasio 80:10:10 untuk Pelatihan, Validasi, dan Pengujian.<sup>1</sup>

- **Hyperparameters:**

- Learning Rate : Nilai kecil ini standar untuk *fine-tuning* agar tidak merusak bobot *pre-trained* yang sudah baik.<sup>16</sup>
- Batch Size : 16 atau 32, disesuaikan dengan memori GPU.
- Optimizer : AdamW (*Adaptive Moment Estimation with Weight Decay*).
- Epochs:
  - NER : 3 Epochs. Dataset NER biasanya konvergen dengan cepat.<sup>1</sup>
  - Sentiment : 2 Epochs. Risiko *overfitting* tinggi jika dilatih terlalu lama pada dataset kecil.<sup>1</sup>
- **Monitoring:** Melacak *Training Loss* dan *Validation Loss* di setiap langkah untuk mendeteksi *overfitting* dini.

## BAB V

### 5. EVALUASI

#### 5.1. Metrik Evaluasi

Evaluasi kinerja model NLP memerlukan metrik yang lebih nuansa daripada sekadar akurasi sederhana, terutama untuk dataset yang tidak seimbang.

1. Untuk NER: Evaluasi dilakukan pada tingkat entitas (*span-level*), bukan token.
  - Precision : Dari semua entitas yang diprediksi model, berapa persen yang benar? (Penting untuk menghindari *False Positives*).
  - Recall : Dari semua entitas yang ada di data nyata, berapa persen yang berhasil ditemukan model? (Penting untuk menghindari *False Negatives*).
  - F1-Score : Rata-rata harmonis dari Precision dan Recall. Ini adalah metrik utama untuk NER.<sup>1</sup>

*Catatan:* Akurasi token diabaikan karena dominasi label 'O' yang dapat memberikan skor tinggi palsu.
2. Untuk Analisis Sentimen:
  - Accuracy: Persentase prediksi yang benar secara keseluruhan.
  - F1-Score: Menilai keseimbangan performa antara kelas positif dan negatif.<sup>1</sup>

#### 5.2. Hasil Eksperimen

Berikut adalah hasil evaluasi kuantitatif dari proses pelatihan yang dilakukan.

**Tabel 5.1: Kinerja Pelatihan Model NER (DistilBERT)**

Epoch	Accuracy (Token)	Precision	Recall	F1-Score	Training Loss	Validation Loss
1	95.41%	73.50%	79.16%	76.23%	0.3935	0.3374
2	95.82%	77.09%	80.29%	78.66%	0.3035	0.2952

3	95.97%	77.71%	81.34%	79.48%	0.2904	0.2375
<b>Final Test</b>	-	<b>79.12%</b>	<b>82.18%</b>	<b>80.62%</b>	<b>0.2807</b>	-

#### Analisis Hasil NER:

Model NER mencapai F1-Score akhir 80.62%. Nilai Recall (82.18%) yang lebih tinggi daripada Precision (79.12%) menunjukkan kecenderungan model yang "agresif"; model lebih memilih untuk menandai sesuatu sebagai entitas daripada melewatkannya. Dalam konteks bisnis restoran, ini seringkali lebih diinginkan (lebih baik sistem menandai "air putih" sebagai item menu yang perlu diperhatikan daripada melewatkannya keluhan pelanggan tentangnya). Penurunan loss yang konsisten menandakan proses belajar yang stabil tanpa tanda-tanda overfitting yang parah.

**Tabel 5.2: Kinerja Pelatihan Model Analisis Sentimen**

Epoch	Accuracy	Precision	Recall	F1-Score	Training Loss	Validation Loss
1	93.0%	93.01%	93.0%	92.99%	0.2899	0.2119
2	93.0%	93.01%	93.0%	92.99%	0.0780	0.1997
<b>Final Test</b>	<b>93.0%</b>	<b>93.02%</b>	<b>93.0%</b>	<b>93.0%</b>	<b>0.2474</b>	-

#### Analisis Hasil Sentimen:

Model sentimen menunjukkan performa yang sangat kuat dengan akurasi 93.0% hanya dalam 2 epoch. Penurunan drastis pada training loss (0.28 ke 0.07) menunjukkan kapasitas DistilBERT untuk beradaptasi dengan cepat pada tugas klasifikasi biner teks ulasan. Performa yang stabil antara validasi dan tes menunjukkan model memiliki kemampuan generalisasi yang baik terhadap data yang belum pernah dilihat sebelumnya.

### 5.3. Analisis Kualitatif: Studi Kasus

Untuk membuktikan efektivitas sistem gabungan (SA-NER), kami menguji model pada kalimat yang kompleks:

"The sushi was fresh, but the waiter was rude."

Proses analisis sistem:

1. **Segmentasi:** Kalimat dipisah berdasarkan tanda baca koma dan konjungsi "but".
  - o Segmen 1: "The sushi was fresh"
  - o Segmen 2: "but the waiter was rude"
2. **NER:**
  - o Segmen 1: "sushi" terdeteksi sebagai Dish.
  - o Segmen 2: "waiter" terdeteksi sebagai Amenity/Service.
3. **Sentimen:**
  - o Segmen 1 diklasifikasikan Positive (probabilitas > 0.9).
  - o Segmen 2 diklasifikasikan Negative (probabilitas > 0.8).

**Tabel 5.3: Hasil Ekstraksi Entitas & Sentimen**

Segmen	Entitas Terdeteksi	Prediksi Sentimen	Interpretasi Bisnis
<i>The sushi was fresh</i>	<b>sushi</b>	<b>Positive</b>	Kualitas produk baik. Pertahankan standar dapur.
<i>The sushi was fresh</i>	<b>fresh</b>	<b>Positive</b>	-
<i>but the waiter was rude</i>	<b>waiter</b>	<b>Negative</b>	Kualitas layanan buruk. Perlu pelatihan staf.
<i>but the waiter was rude</i>	<b>rude</b>	<b>Negative</b>	-

Hasil ini (Tabel 5.3) menegaskan bahwa sistem mampu memisahkan sentimen berdasarkan aspeknya. Jika menggunakan analisis sentimen standar pada seluruh kalimat, hasilnya kemungkinan besar "Netral" (karena kata positif dan negatif saling meniadakan) atau salah satu sentimen mendominasi, menghilangkan informasi krusial bagi manajemen restoran.<sup>1</sup>

## BAB VI

### 6. DEPLOYMENT

#### 6.1. Arsitektur Deployment dengan Streamlit

Untuk membuat model dapat diakses oleh pengguna akhir (pemilik restoran atau manajer), kami mengembangkan aplikasi web menggunakan **Streamlit**. Streamlit adalah kerangka kerja Python *open-source* yang memungkinkan pembuatan aplikasi data interaktif dengan cepat tanpa memerlukan keahlian pemrograman *frontend* (HTML/CSS) yang mendalam.<sup>17</sup>

Arsitektur aplikasi terdiri dari komponen berikut:

1. **Model Loader:** Menggunakan dekorator `@st.cache_resource` dari Streamlit. Ini sangat penting untuk memuat model DistilBERT yang berukuran besar (ratusan MB) ke dalam memori hanya satu kali saat aplikasi pertama kali dijalankan, bukan setiap kali pengguna melakukan interaksi. Ini menjaga latensi tetap rendah.<sup>18</sup>
2. **Input Interface:** Area teks (`st.text_area`) bagi pengguna untuk menempelkan ulasan restoran.
3. **Processing Engine:** Backend Python yang menjalankan fungsi tokenisasi, inferensi model NER, dan inferensi model Sentimen secara berurutan.
4. **Visualization Module:** Menggunakan pustaka visualisasi seperti `spacy-streamlit` atau komponen HTML kustom untuk merender teks dengan entitas yang diwarnai (misalnya, warna hijau untuk entitas bersentimen positif dan merah untuk negatif).<sup>19</sup>

## 6.2. Fitur Aplikasi

Aplikasi yang dibangun memiliki fitur utama:

- **Analisis Real-time:** Pengguna mendapatkan hasil analisis instan setelah menekan tombol "Analyze".
- **Highlighting Entitas:** Entitas dalam teks diberi sorotan warna sesuai dengan kategori (Makanan, Lokasi, Harga).
- **Indikator Sentimen:** Ikon emoji atau label warna (Hijau/Merah) ditampilkan di sebelah entitas untuk menunjukkan sentimen.
- **Tabel Ringkasan:** Menyajikan daftar semua entitas yang ditemukan beserta skor kepercayaan (*confidence score*) model.

## BAB VII

### 7. REFLEKSI & DISKUSI

#### 7.1. Pembahasan Temuan

Penelitian ini memvalidasi hipotesis bahwa model bahasa yang lebih kecil dan efisien seperti DistilBERT sangat kapabel untuk menangani tugas NLP yang kompleks di domain spesifik seperti ulasan restoran. Dengan F1-Score ~80% untuk NER dan Akurasi ~93% untuk Sentimen, sistem ini sudah mencapai tingkat kelayakan untuk penggunaan komersial dasar. Temuan kunci lainnya adalah pentingnya konteks lokal. Model terbukti mampu membedakan kata yang sama dengan makna berbeda. Misalnya, kata "hot" dalam "hot soup" (positif, terdeteksi sebagai deskripsi Dish) vs "hot room" (negatif, terdeteksi sebagai deskripsi Amenity). Hal ini tidak mungkin dicapai dengan metode *keyword matching* sederhana.

#### 7.2. Limitasi Sistem

Meskipun hasilnya menjanjikan, beberapa keterbatasan teridentifikasi selama proses evaluasi:

1. **Ketergantungan pada Segmentasi Eksplisit:** Sistem saat ini memecah kalimat

berdasarkan tanda baca (koma, titik). Jika pengguna menulis ulasan panjang tanpa tanda baca ("makanannya enak tempatnya kotor pelayannya lambat"), segmentasi akan gagal, dan sentimen dari satu bagian mungkin "bocor" ke entitas di bagian lain.

2. **Dataset Domain Terbatas:** Model NER dilatih pada *MIT Corpus* yang lebih berorientasi pada pertanyaan pencarian (*queries*), bukan ulasan naratif panjang. Hal ini kadang menyebabkan model kesulitan mengenali entitas dalam kalimat yang sangat panjang dan berbunga-bunga.
3. **Aspek Implisit:** Model hanya mengenali entitas yang tertulis secara eksplisit. Jika ulasan berbunyi "Harganya terlalu mahal", model mengenali "Harga". Namun, jika ulasan berbunyi "Dompet saya menangis", model mungkin gagal mengasosiasikan ini dengan aspek "Harga" karena tidak ada entitas eksplisit yang sesuai dengan pelatihan NER.
4. **Bahasa:** Model saat ini bersifat monolingual (Bahasa Inggris). Tidak dapat menangani ulasan Bahasa Indonesia atau *code-mixing* (campur kode) yang sangat umum di Indonesia (misal: "Taste-nya okay tapi price-nya mahal").

### 7.3. Refleksi

Penggunaan AI untuk analisis sentimen harus mempertimbangkan privasi. Meskipun ulasan bersifat publik, sistem tidak boleh digunakan untuk memprofilkan individu pengulas secara spesifik, melainkan hanya untuk agregasi tren kualitas layanan.

## BAB VIII

### 8. KESIMPULAN & SARAN

#### 8.1. Kesimpulan

Proyek ini telah berhasil mengembangkan sistem *Sentiment-Aware Named Entity Recognition* (SA-NER) menggunakan model DistilBERT yang di-*fine-tune*. Sistem ini secara efektif mengatasi masalah informasi yang berlebihan (*information overload*) yang dihadapi pemilik restoran dengan mengotomatisasi ekstraksi wawasan dari ulasan pelanggan.

Secara teknis, integrasi dua model terpisah (NER dan Sentimen) terbukti menjadi strategi yang efektif, memberikan fleksibilitas dan akurasi tinggi (NER F1: 80.62%, Sentimen Acc: 93.0%). Penggunaan DistilBERT juga terbukti tepat, memberikan keseimbangan antara performa model dan efisiensi sumber daya, yang memungkinkan *deployment* aplikasi yang responsif pada platform Streamlit. Sistem ini menawarkan solusi konkret untuk meningkatkan pengambilan keputusan berbasis data dalam manajemen kualitas restoran.

#### 8.2. Saran

Untuk pengembangan selanjutnya, disarankan untuk:

1. Menangani Aspek Implisit: Mengembangkan modul tambahan untuk *Implicit Aspect Extraction* guna menangkap keluhan yang tidak menyebutkan nama benda secara langsung.
2. Model Multibahasa: Mengganti DistilBERT dengan mBERT (Multilingual BERT) atau

XLM-RoBERTa agar aplikasi dapat memproses ulasan dalam Bahasa Indonesia dan Bahasa Inggris sekaligus, meningkatkan relevansi pasar lokal.

3. Segmentasi Berbasis Dependensi: Menggunakan *Dependency Parsing* (seperti pustaka spaCy) untuk menghubungkan kata sifat sentimen ke target entitasnya secara gramatis, alih-alih hanya memotong kalimat berdasarkan koma. Ini akan meningkatkan akurasi pada kalimat tanpa tanda baca.
4. Perluasan Dataset: Menambah dataset pelatihan dengan ulasan naratif asli dari Google Maps/Zomato untuk memperkaya variasi konteks yang dipelajari model NER.

## REFERENSI

1. P. R. Amalia, E. Winarko, Aspect-based sentiment analysis on indonesian restaurant review using a combination of convolutional neural network and contextualized word embedding, *Indonesian Journal of Computing and Cybernetics Systems* 15 (3) (2021) 285–294. doi: 10.22146/ijccs.67306
2. Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. <https://www.emergentmind.com/topics/distilbert-base>
3. Towards Data Science. (2024). *A Complete Guide to BERT with Code*. <https://towardsdatascience.com/a-complete-guide-to-bert-with-code-9f87602e4a11/>
4. Machine Learning Mastery. (2024). *How to do Named Entity Recognition (NER) with a BERT model*. <https://machinelearningmastery.com/how-to-do-named-entity-recognition-ner-with-a-bert-model/>
5. Introduction to DistilBERT Model - GeeksforGeeks <https://www.geeksforgeeks.org/nlp/introduction-to-distilbert-model/>
6. DistilBERT in Natural Language Processing - GeeksforGeeks <https://www.geeksforgeeks.org/nlp/distilbert-in-natural-language-processing/>
7. Wikipedia. (2024). *Inside–outside–beginning (tagging)*. [https://en.wikipedia.org/wiki/Inside%E2%80%93outside%E2%80%93beginning\\_\(tagging\)](https://en.wikipedia.org/wiki/Inside%E2%80%93outside%E2%80%93beginning_(tagging))
8. Kanerika. (2024). *Named Entity Recognition: A Comprehensive Guide to NLP's Key Technology*. <https://medium.com/@kanerika/named-entity-recognition-a-comprehensive-guide-to-nlps-key-technology-636a124eaa46>
9. STEMM Institute Press. (2024). *Comparison of DistilBERT and LLaMA 3 for Sentiment Analysis*. Journal of Big Data and Computing. <http://www.stemmpress.com/uploadfile/202501/3d5320845cb2692.pdf>
10. tner/mit\_restaurant . Datasets at Hugging Face, [https://huggingface.co/datasets/tner/mit\\_restaurant](https://huggingface.co/datasets/tner/mit_restaurant)
11. Signature redacted - DSpace@MIT, <https://dspace.mit.edu/bitstream/handle/1721.1/122557/1122790819-MIT.pdf?sequence=1&isAllowed=y>

12. Maru, S. (2024). MIT Restaurant Review - CRF (NER) Model. Kaggle. <https://www.kaggle.com/code/marusagar/mit-restaurant-review-crf-ner-model>
13. Restaurant Reviews for Sentiment Analysis - Kaggle, <https://www.kaggle.com/datasets/ziadmostafa1/restaurant-reviews>
14. Sentiment Analysis of Restaurant Reviews | by Vincents - Medium, <https://medium.com/@vincents0511/sentiment-analysis-of-restaurant-reviews-430875614095>
15. Towards Data Science. (2024). *Fine-tuning DistilBERT for Binary Classification Tasks.* <https://towardsdatascience.com/hugging-face-transformers-fine-tuning-distilbert-for-binary-classification-tasks-490f1d192379/>
16. amanpatkar/distilbert-finetuned-ner - Hugging Face, <https://huggingface.co/amanpatkar/distilbert-finetuned-ner>
17. Streamlit Inc. (2024). *Streamlit: The fastest way to build and share data apps.*
18. Build a Sentiment Analysis App with Hugging Face and Streamlit - Codecademy, <https://www.codecademy.com/article/hugging-face-with-streamlit-app>
19. Components - Streamlit, <https://streamlit.io/components>
20. Visualizers · spaCy Usage Documentation, <https://spacy.io/usage/visualizers>

## LAMPIRAN

### 1. Repository GitHub:

<https://github.com/RainerYesaya/Sentiment-Aware-NER-for-Analyzing-Food-and-Service-Quality-in-Restaurants>

### 2. Dataset:

- MIT Restaurant Corpus: <https://www.kaggle.com/datasets/marusagar/mit-restaurant-corpus-crf-dataset>
- Sentiment Dataset: <https://github.com/aadimangla/Restaurant-Reviews-Sentiment-Analysis>

### 3. Kode Snippet:

- main\_app.ipynb:

```

from transformers import DistilBertForSequenceClassification, DistilBertTokenizerFast, DistilBertForTokenClassification
import torch
import re
import pandas as pd
import gradio as gr
import matplotlib.pyplot as plt

DATASET_PATH = "./dataset"
MODEL_PATH = "./trained_models"

SA_DATA = DATASET_PATH + "/sa"
SA_MODEL = MODEL_PATH + "/sa"

NER_DATA = DATASET_PATH + "/ner"
NER_MODEL = MODEL_PATH + "/ner"

[24] ✓ 0.6s

sentiment_model = DistilBertForSequenceClassification.from_pretrained(SA_MODEL)
sentiment_tokenizer = DistilBertTokenizerFast.from_pretrained(SA_MODEL)

ner_model = DistilBertForTokenClassification.from_pretrained(NER_MODEL)
ner_tokenizer = DistilBertTokenizerFast.from_pretrained(NER_MODEL)

print("Model loaded successfully!")

[25] ✓ 0.6s

... Model loaded successfully!

```

- views:

### Sentiment-Aware Named Entity Recognition for Analyzing Food and Service Quality in Restaurants

Automatic detection of restaurant aspects (Food, Service, Ambience) and their associated sentiment polarity using Deep Learning.

**Background Issue**

In the restaurant industry, customer reviews are a valuable asset. However, manually reading thousands of reviews is time-consuming and prone to bias. Restaurant owners often struggle to know: *'Are customers disappointed because of the food taste or the service?'* This system aims to automatically separate aspects (Entities) and their sentiment polarity.

**Model Explanation**

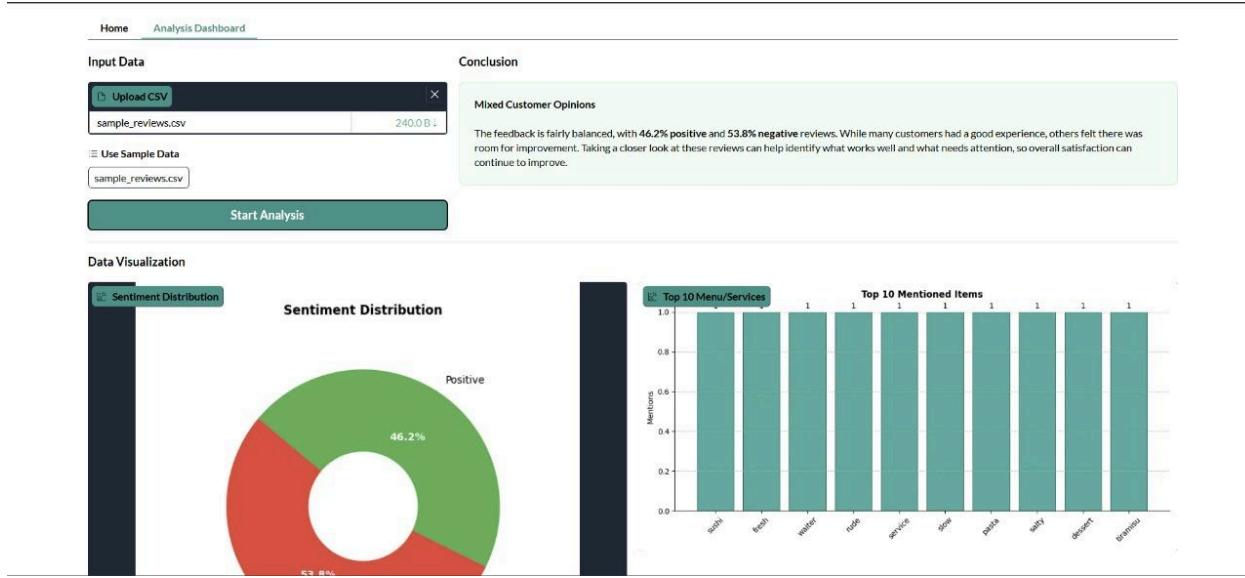
This application uses **Deep Learning (DistilBERT)** which works in two stages:

- **NER (Named Entity Recognition):** Detects keywords such as *Sushi, Pizza* (Food) or *Waiter* (Service).
- **Sentiment Analysis:** Determines whether the sentence has a **Positive** or **Negative** meaning.

**How to Use the Application**

1. Go to the **Analysis Dashboard** tab at the top.
2. Upload a CSV file that has a column named **review**.
3. Click the **Start Analysis** button.
4. Analysis results in the form of Graphs and Conclusions will appear automatically.

Correct CSV Data Format Example:



#### 4. Demo Application:

- [https://www.canva.com/design/DAG7daycpWo/D90eNSFIrU5DgQoCiXuhA/edit?utm\\_content=DAG7daycpWo&utm\\_campaign=designshare&utm\\_medium=link2&utm\\_source=sharebutton](https://www.canva.com/design/DAG7daycpWo/D90eNSFIrU5DgQoCiXuhA/edit?utm_content=DAG7daycpWo&utm_campaign=designshare&utm_medium=link2&utm_source=sharebutton)
- (Backup): [AOLDeepLearning\\_Sentiment-Aware NER for Analyzing Food and Serv...](#)