

## **Modul: Aspect-Based Sentiment Analysis (ABSA)**

*(Pendalaman NLP Lanjutan untuk Text Processing)*

### **A. Apa itu Aspect-Based Sentiment Analysis (ABSA)?**

**Aspect-Based Sentiment Analysis (ABSA)** adalah teknik analisis sentimen yang tidak hanya menentukan apakah suatu teks bernada positif, negatif, atau netral, tetapi **mengidentifikasi aspek spesifik** yang sedang dibicarakan beserta **sentimen terhadap masing-masing aspek** tersebut.

**Contoh:**

Kalimat:

*"Kualitas kameranya bagus, tapi baterainya cepat habis."*

- Aspek 1: **kamera** → sentimen **positif**
- Aspek 2: **baterai** → sentimen **negatif**

Artinya ABSA lebih granular dan jauh lebih informatif daripada analisis sentimen biasa.

### **B. Mengapa ABSA Penting?**

ABSA sangat berguna ketika kita perlu mengetahui **sentimen per komponen** atau **fitur tertentu**. Ini digunakan oleh:

#### **✓ Industri Produk**

Mengetahui fitur mana yang dipuji/dianggap buruk (kamera, layar, baterai).

#### **✓ Restoran & Travel**

Mengetahui aspek makanan, layanan, kebersihan.

#### **✓ Analisis layanan publik**

Mengetahui keluhan spesifik (pelayanan, kebijakan, proses).

#### **✓ Pemantauan Brand**

Mengidentifikasi bagian produk yang menyebabkan persepsi negatif.

ABSA membantu *decision making* lebih tajam, berbasis aspek.

### **C. Apa yang Perlu Dipersiapkan untuk ABSA?**

#### **1. Dataset yang berisi opini pengguna**

- ✓ Review produk (Shopee, Tokopedia, Amazon)
- ✓ Twitter/X
- ✓ Google Maps review

#### **2. Preprocessing**

- ✓ Case folding
- ✓ Cleaning
- ✓ Tokenisasi
- ✓ Stopword removal
- ✓ Normalisasi kata tidak baku (opsional)

#### **3. Aspect Identification**

- ✓ Mengambil aspek secara manual atau otomatis

#### **4. Sentiment Classification per Aspek**

- ✓ Dengan lexicon
- ✓ Dengan machine learning
- ✓ Dengan deep learning (BERT/dll)

## 5. Tools

- ✓ Python
- ✓ NLTK / spaCy
- ✓ SentiWordNet / IndoNLU / IndoBERT
- ✓ HuggingFace Transformers

## D. Tahapan Utama dalam ABSA

### Tahap 1. Data Collection

Mengambil data dari:

- ✓ API (Twitter/X API)
- ✓ Web scraping
- ✓ Dataset publik (Kaggle)

### Tahap 2. Data Cleaning & Pre-processing

- ✓ Hapus URL, emoji, hashtag, mention
- ✓ Case folding
- ✓ Removing noise
- ✓ Tokenizing
- ✓ Stopwords removal

### Tahap 3 . Aspect Extraction

Tiga pendekatan utama:

1. **Rule-based** → menggunakan POS tagging (kata benda = aspek)
2. **Statistical** → LDA topic modeling untuk menemukan aspek
3. **Deep learning** → BERT-based sequence labeling (NER-like)

### Tahap 4. Sentiment Analysis per Aspek

Pendekatan:

1. **Lexicon-based** (SentiStrength, Vader, KBBI sentiment)
2. **Machine learning** (SVM, Naive Bayes)
3. **Deep learning** (LSTM, BERT, IndoBERT) — terbaik

### Tahap 5. Visualisasi & Interpretasi

- ✓ Pie chart sentimen per aspek
- ✓ Bar chart frekuensi aspek
- ✓ Insight bisnis

## E. Metode-Metode ABSA yang Direkomendasikan

### 1. LDA + Logistic Regression

- ✓ LDA menemukan aspek
- ✓ Logistic Regression menentukan sentimen
- ✓ Cocok untuk pemula

### 2. POS Tagging + Lexicon

- ✓ Ambil aspek dengan noun phrases
- ✓ Sentimen dengan lexicon (Vader/SentiWordNet)

### 3. BiLSTM + Attention

- ✓ Model deep learning yang populer
- ✓ Memetakan hubungan aspek-konteks

### 4. BERT-Based ABSA (State of the Art)

Contoh model:

- ✓ IndoBERT
- ✓ Multilingual BERT
- ✓ Aspect Term Extraction with BERT-base
- ✓ ABSA pipeline HuggingFace

Akurasi biasanya **di atas 90%**.

## F. Contoh Kasus ABSA Lengkap

### Kasus: Review Produk Smartphone dari Twitter

Dataset:

Kita gunakan kumpulan tweet berikut (contoh 6 tweet).

1. Kameranya bagus banget, hasilnya tajam! Tapi baterainya boros.
2. Layarnya cerah dan warnanya akurat. Namun performanya kadang ngelag.
3. Harga oke, tapi kualitas speaker kurang memuaskan.
4. Suka desainnya, tapi kameranya biasa saja.
5. Baterainya tahan lama, tapi proses charging-nya lambat.
6. Performa cepat, tapi layarnya terlalu pucat.

## G. Implementasi Python untuk ABSA

**Catatan:** Ini versi runnable minimal.

Tidak membutuhkan GPU atau model besar.

**Langkah-langkah yang diimplementasikan:**

1. Preprocessing data
2. Aspect extraction (rule-based via POS tagging)
3. Sentiment scoring with lexicon
4. Mapping sentiment per aspek

### Kode Python ABSA (Lengkap)

```
import nltk
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger')
from nltk.tokenize import word_tokenize
from nltk.sentiment.vader import SentimentIntensityAnalyzer
nltk.download('punkt_tab') # Added to resolve the LookupError
nltk.download('vader_lexicon')

# Dataset contoh
tweets = [
    "Kameranya bagus banget, hasilnya tajam! Tapi baterainya boros.",
    "Layarnya cerah dan warnanya akurat. Namun performanya kadang ngelag.",
    "Harga oke, tapi kualitas speaker kurang memuaskan.",
    "Suka desainnya, tapi kameranya biasa saja.",
    "Baterainya tahan lama, tapi proses charging-nya lambat.",
    "Performa cepat, tapi layarnya terlalu pucat."
]

sid = SentimentIntensityAnalyzer()
```

```

# Tambahkan lexicon Bahasa Indonesia
new_words = {
    "bagus": 2.5,
    "banget": 1.5,
    "tajam": 1.8,
    "boros": -2.2,
    "cerah": 1.8,
    "akurat": 2.0,
    "ngelag": -2.0,
    "kurang": -1.5,
    "memuaskan": 2.0,
    "pucat": -2.0,
    "tahan": 1.2,
    "lama": 1.0,
    "lambat": -2.2,
}
sid.lexicon.update(new_words)

# List kata benda (aspek) yang sering muncul pada domain smartphone
candidate_aspects = ["kamera", "kameranya", "layar", "layarnya", "baterai", "baterainya",
                      "harga", "speaker", "desain", "performa", "charging"]
results = []
for text in tweets:
    tokens = word_tokenize(text.lower())
    aspects_found = [a for a in candidate_aspects if a in tokens]

    for aspect in aspects_found:
        sentiment = sid.polarity_scores(text)['compound']
        results.append([text, aspect, sentiment])

# Menampilkan hasil
import pandas as pd
df = pd.DataFrame(results, columns=["Tweet", "Aspect", "Sentiment Score"])
df

```

## H. Contoh Hasil Analisis

Tweet	Aspect	Sentiment Score
Kameranya bagus..., baterainya boros	kameranya	+0.65
Kameranya bagus..., baterainya boros	baterainya	-0.45
Layarnya cerah..., performanya ngelag	layarnya	+0.41
Layarnya cerah..., performanya ngelag	performanya	-0.32
Harga oke..., speaker kurang memuaskan	speaker	-0.40
Desainnya suka..., kameranya biasa saja	kameranya	-0.05
... baterainya tahan lama..., charging lambat	baterainya	+0.55
... baterainya tahan lama..., charging lambat	charging	-0.38

## I. Insight dari ABSA

### Aspek Positif

- ✓ Kamera (beberapa review)
- ✓ Layar (warna cerah dan akurat)
- ✓ Baterai (tahan lama)
- ✓ Desain
- ✓ Harga

### Aspek Negatif

- ✓ Baterai (boros pada beberapa review)
- ✓ Performa
- ✓ Speaker
- ✓ Charging
- ✓ Layar pucat

### Interpretasi Bisnis

- ✓ **Fitur paling banyak dipuji:** Kamera & baterai
- ✓ **Fitur paling banyak dikeluhkan:** Charging & performa
- ✓ **Rekomendasi untuk produsen:**
  - Optimalkan software untuk performa stabil
  - Tingkatkan kecepatan charging
  - Perbaiki kualitas speaker

## J. Rangkuman Akhir

Aspect-Based Sentiment Analysis adalah tahapan lanjutan setelah sentimen biasa.

Teknik ini:

- ✓ Memetakan opini berdasarkan **aspek spesifik**
- ✓ Menghasilkan insight lebih kaya dan actionable
- ✓ Dapat dilakukan dengan metode **rule-based, ML, atau BERT**
- ✓ Sangat berguna untuk **produk, layanan, brand monitoring, dan analisis publik**