

LEMBAR LATIHAN: Part-of-Speech Tagging (POS Tagging)

Mata Kuliah: Natural Language Processing (NLP)

Topik: Open Class dan Closed Class pada POS Tagging

Tujuan Pembelajaran

- ✓ Menjelaskan konsep dasar POS Tagging.
- ✓ Mengidentifikasi kelas kata dalam kalimat.
- ✓ Membedakan antara kelas kata Open Class dan Closed Class.
- ✓ Menerapkan POS Tagging menggunakan Python.

A. Teori Singkat (Isian)

1. Part of Speech (POS) merupakan klasifikasi dari kata-kata berdasarkan _____ dan _____ dalam struktur kalimat.
2. Contoh noun adalah _____, sedangkan contoh verb adalah _____.
3. Kelompok kata yang termasuk open class adalah _____, sedangkan kelompok closed class adalah _____.
4. Tujuan utama dari POS Tagging adalah _____.
5. Dua tantangan utama dalam POS Tagging adalah _____ dan _____.
6. Pendekatan POS Tagging dapat dilakukan dengan metode _____ (berbasis aturan) dan _____ (berbasis pembelajaran).

B. Latihan Pilihan Ganda

1. Kelas kata yang jumlahnya tetap dan jarang berubah termasuk ...
 - a. Verb
 - b. Adjective
 - c. Closed Class
 - d. Open Class
2. Kata "menyanyi" termasuk kategori ...
 - a. Noun
 - b. Verb (Open Class)
 - c. Conjunction
 - d. Preposition
3. POS Tagging diperlukan dalam sistem seperti berikut, kecuali ...
 - a. Speech synthesis
 - b. Machine translation
 - c. Parsing
 - d. Image segmentation
4. Tantangan utama POS Tagging yang berkaitan dengan "kata yang belum pernah ditemui" adalah ...
 - a. Sparse data
 - b. Ambiguity
 - c. Noise data
 - d. Overfitting
5. Dalam POS Tagging, model learning-based menggunakan pendekatan ...
 - a. Rule manual
 - b. Statistik atau Machine Learning
 - c. Kamus tetap
 - d. Ontologi

C. Identifikasi Kelas Kata

No	Kata	Kelas POS	Kategori (Open/Closed)
1	berjalan
2	saya
3	cantik
4	di
5	mahasiswa
6	atau
7	melalui
8	bawah

D. Analisis Singkat

1. Jelaskan perbedaan antara open class dan closed class dalam POS Tagging.
2. Mengapa kata seperti 'dan' atau 'atau' termasuk closed class?
3. Sebutkan contoh ambiguitas dalam POS Tagging yang dapat menyebabkan kesalahan label.
4. Jelaskan bagaimana learning-based method membantu mengatasi masalah sparse data.

E. Latihan Pemrograman

Gunakan Python untuk melakukan POS Tagging sederhana.

Soal 1: Tagging Dasar

Tulislah kode Python untuk melakukan POS Tagging pada kalimat:
"Mahasiswa sedang belajar NLP di laboratorium komputer."

```
import nltk
from nltk import word_tokenize, pos_tag
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger')

kalimat = "Mahasiswa sedang belajar NLP di laboratorium komputer."
tokens = word_tokenize(kalimat)
tagging = pos_tag(tokens)
print(tagging)
```

Tugas:

1. Jalankan kode di atas.
2. Salin hasil tagging.
3. Identifikasi mana yang termasuk open class dan closed class.

Soal 2: Menggunakan Tagset Universal

Tambahkan argumen untuk menggunakan tagset universal:

```
pos_tag(tokens, tagset='universal')
```

Pertanyaan:

1. Apa perbedaan hasil tagging dengan tagset universal?
2. Bagaimana Anda menginterpretasikan tag ADJ, ADV, PRON, ADP?

Soal 3: Analisis Hasil

1. Coba ubah kalimat menjadi: "Dia suka minum air putih di pagi hari." Apakah ada tag yang ambigu? Jelaskan.
2. Menurut Anda, tantangan apa yang muncul jika kalimat mengandung kata tidak baku seperti "ngopi" atau "nongkrong"?

---00---