

Laporan Analisis Word Sense Disambiguation

1. Buatlah program untuk menghitung skor kemiripan/similarity antara 2 buah teks dengan algoritma Lesk, yaitu berdasarkan jumlah kata yang beririsan pada 2 buah teks tersebut

```
compute_overlap('Aziz mempunyai mental juara','Bola itu mental terkena tiang gawang')
```

```
69
```

```
compute_overlap('ditemukan mayat tanpa kepala','saya membuat kepala surat')
```

```
73
```

2. Pilihlah tiga contoh kata ambigu, kemudian buat contoh kalimat yang mengandung kata tersebut. Jadi Anda akan memiliki 6 (3 x 2 kalimat), karena masing-masing sense kata diwakili oleh 1 kalimat.

```
simplified_lesk('bisa','ia bisa membaca, tetapi tidak bisa menulis')
```

```
'bi.sa1'
```

```
[48] simplified_lesk('bisa','saya terkena bisa ular')
```

```
'bi.sa2'
```

```
[49] simplified_lesk('kepala','ditemukan mayat tanpa kepala')
```

```
'ke.pa.la1'
```

```
[63] simplified_lesk('kepala','saya membuat kepala surat')
```

```
'ke.pa.la2'
```

```
[52] simplified_lesk('mental','Bola itu mental terkena tiang gawang.')
```

```
'men.tal1'
```

```
[56] simplified_lesk('mental','Aziz mempunyai mental juara')
```

```
'men.tal2'
```

3 Alasan pemilihan kata-kata ambigu

Pemilihan kata berdasarkan riset yaitu tentang homograf yang berasal dari kata *homo* yang berarti ‘sama’ dan *graf* atau *graph* yang berarti ‘tulisan’. Sehingga secara etimologi, istilah homograf dapat diartikan sebagai kata atau istilah yang memiliki cara penulisan atau ejaan sama namun berbeda pengucapan atau pelafalannya. Sama seperti istilah homofon, istilah –istilah atau kata – kata homograf tidak memiliki kesamaan arti satu sama lain. KBBI juga memberikan pengertian terhadap istilah homograf sebagai kata yang sama ejaannya dengan kata lain, tetapi berbeda lafal dan maknanya (misal teras ‘inti kayu’ dengan teras ‘bagian rumah’). Namun, berbeda dengan homonim dan homofon yang cenderung susah diidentifikasi, homograf lebih mudah diidentifikasi meskipun tidak digabungkan dalam suatu kalimat utuh. Berdasarkan kata yang sama ejaannya dengan kata lain tersebut menjadi pemilihan kata-kata “Bisa”, “Mental” dan juga “Kepala” karena dengan ejaan yang sama, tetapi menghasilkan lafal dan makna yang berbeda

4. Analisis terhadap hasil identifikasi sense yang tepat pada keenam contoh kalimat yang disebutkan pada nomor 2

- Bisa

Analisis kata “bisa” pada kalimat uji “ia bisa membaca, tetapi tidak bisa menulis” menghasilkan sense bi.sa1 sangat tepat karena pada sense bi.sa1 kata bisa memiliki makna mampu membaca, selanjutnya kata “bisa” pada kalimat uji “saya terkena bisa ular” menghasilkan sense bi.sa2 sangat tepat karena pada sense bi.sa2 kata bisa memiliki makna zat racun dimana bisa ular memiliki racun.

-Kepala

Analisis kata “kepala” pada kalimat uji “ditemukan mayat tanpa kepala” menghasilkan sense ke.pa.la1 sangat tepat karena pada sense ke.pa.la1 kata kepala memiliki makna bagian tubuh yang di atas leher, selanjutnya kata “kepala” pada kalimat uji “saya membuat kepala surat” menghasilkan sense ke.pa.la2 sangat tepat karena pada sense ke.pa.la2 kata kepala memiliki makna bagian pengodean yang mendahului tubuh dokumen dan digunakan untuk menggambarkan dokumen itu sendiri karena kepala surat merupakan tubuh dokumen.

-Mental

Analisis kata “Mental” pada kalimat uji “Bola itu mental terkena tiang gawang” menghasilkan sense men.tal1 sangat tepat karena pada sense men.tal1 memiliki makna terpelanting, terpelanting dari tiang gawang, selanjutnya kata “Mental” pada kalimat uji “Aziz mempunyai mental juara” menghasilkan sense men.tal2 sangat tepat karena pada sense men.tal2 memiliki makna ada gunanya, ada pengaruhnya karena mental juara sangat berpengaruh pada batin dan watak manusia.