

Laporan Analisis Tugas Mesin Penerjemah (Irawansyah – 1301174689)

1. Pendahuluan

Machine Translation atau Mesin Penerjemah adalah proses mengubah teks dalam sebuah Bahasa ke Bahasa Lain, pada tugas kali ini saya menggunakan bahasa indonesia sebagai bahasa sumber dan bahasa sunda sebagai bahasa daerah. Contoh seperti gambar berikut:



2. Lakukan pengukuran skor probability alignment 3 pasang kata

```
print(round(com_ibm1.translation_table["tidur"]["bobo"], 3) )
```

0.851

```
print(round(com_ibm1.translation_table["makan"]["emam"], 3) )
```

0.908

```
print(round(com_ibm1.translation_table["kami"]["urang"], 3) )
```

0.767

Berdasarkan skor yang didapat tinggi untuk kata sumber tidur dan kata target bobo mencapai 85% artinya sangat baik artinya bahasa sumber dan target sangat mirip sehingga memiliki arti yang sama, untuk kata sumber makan dan kata target emam mencapai 90% sangat baik artinya bahasa sumber dan target memiliki arti yang sama dan untuk kata sumber kami dan kata target urang mencapai 76% saja sudah baik artinya kata sumber dan target memiliki arti yang sama

3. Lakukan analisis terhadap hasil pengukuran skor probability alignment

a. Perbandingan skor probability, mana yang lebih tinggi, mana yang lebih rendah, mengapa?

Perbandingan kata sumber tidur dan kata target bobo memiliki skor probability yang tinggi dan kata sumber kami dan kata target urang memiliki skor probability yang paling rendah, ini terjadi karena dengan melakukan pendekatan melalui align kalimat yang berdasarkan panjang kata dan karakter. Dengan mengambil irisan alignment sumber-ke-target dan target-ke-sumber, lalu ambil gabungan alignment sumber-ke-target dan target-ke-sumber ini dilakukan secara incremental artinya dilakukan secara berulang dari gabungan ke irisan dan berdasarkan hasil *alignment* terakhir dengan melihat semua pasangan pada frase dimana semua kata didalam align yang sama, jika kata frase terdapat kesamaan di-align maka kata tersebut akan menghasilkan probability yang tinggi.

b. Penjelasan singkat kaitannya dengan metode IBMModel1

Untuk menghasilkan string f bahasa sunda dari bahasa Indonesia string e dengan menggunakan 2 menghitung probability, perhitungan probability dengan menggunakan 2 langkah:

1. Langkah 1 : Pilih alignment alpha dengan probabilitas $\frac{1}{(l+1)^m}$

2. Langkah 2 : Pilih kata kata sunda dengan probabilitas paling tinggi berdasarkan persamaan dibawah ini:

$$p(f|a, e, m) = \prod_{j=1}^m t(f_j|e_{a_j})$$