

LAPORAN TUGAS NLP PEKAN 14 – SISTEM TANYA JAWAB

1. Berikan analisis dugaan Anda mengapa tipe question ABBR kinerjanya paling rendah (sulit dideteksi), sementara tipe question LOC mempunyai kinerja paling tinggi?

Karena tipe pertanyaan ABBR atau abbreviation merupakan pertanyaan yang menanyakan singkatan dari sebuah kata yang panjang atau kepanjangan dari sebuah singkatan yang mana pada setiap bahasa bisa saja kata singkatan tersebut memiliki arti yang berbeda-beda.

Sedangkan, jika tipe pertanyaan LOC atau keberadaan suatu tempat, setiap pertanyaan akan memiliki jawaban yang sama di bahasa manapun karena suatu tempat adalah hal yang pasti.

2. Apakah menurut Anda dapat diterapkan metode gabungan dengan pendefinisian aturan secara manual untuk mendeteksi satu atau beberapa tipe question? Sebutkan contoh aturan yang dapat diterapkan untuk mendeteksi tipe question (dalam bentuk pseudocode/algoritma).

Sebenarnya hal tersebut bisa saja diterapkan, misalkan tipe pertanyaan dimana atau “LOC” dengan pertanyaan “where” digabungkan dengan tipe pertanyaan dengan tipe “NUM” dengan pertanyaan “how many”. Contoh pertanyaannya adalah jika ada yang bertanya tentang berapa jumlah korban dan dirawat di rumah sakit mana. Cara mendeteksi hal tersebut adalah dengan membuat aturan dalam bentuk pseudocode atau algoritma yang seperti ini :

```
elif (((“How much” or “How many”) and “Where”) in current_question):  
    current_label = “NUM” and “LOC”
```

Maka tipe question tersebut akan mendapatkan dua label.

3. Terapkan aturan tambahan yang didefinisikan secara manual tersebut untuk mengidentifikasi tipe question pada dataset uji yang sama, lalu amati hasil evaluasinya!

Ditambah dengan aturan-aturan yang baru, seperti :

- Jika pertanyaannya berupa “Who” atau “Whom” atau “Whose” maka otomatis pertanyaan tersebut akan diberi label “HUM”
- Jika pertanyaannya berupa “How much” atau “How many” maka otomatis pertanyaan tersebut akan diberi label “NUM”
- Jika pertanyaannya berupa “abbreviation” atau “stands for” maka otomatis pertanyaan tersebut akan diberi label “ABBR”
- Jika pertanyaannya berupa “Where” maka otomatis pertanyaan tersebut akan diberi label “LOC”

```
labels = []
questions = []
for line in lines:
    tokens = line.split()
    current_label = tokens[0].split(':')[0]
    current_question = line[len(tokens[0])+1:]
    if (("Who" or "Whom" or "Whose") in current_question):
        current_label = "HUM"
    elif (("How much" or "How many") in current_question):
        current_label = "NUM"
    elif (("abbreviation" or "stands for") in current_question):
        current_label = "ABBR"
    elif ("Where" in current_question):
        current_label = "LOC"

    labels.append(current_label)
    # 'Path=helloworld'.removeprefix('Path=')
    # current_question = line[len(tokens[0])+1:]
    questions.append(current_question)
```

Setelah ditambahkan aturan baru secara manual seperti algoritma di atas, maka yang berubah adalah nilai dari precision, recall, f1-score dan akurasi. Mari kita bandingkan.

Nilai akurasi sebelum ditambahkan aturan baru :

	precision	recall	f1-score	support
ABBR	0.00	0.00	0.00	4
DESC	0.71	0.54	0.62	37
ENTY	0.52	0.67	0.59	51
HUM	0.78	0.76	0.77	50
LOC	0.85	0.88	0.86	32
NUM	0.80	0.77	0.78	26
accuracy			0.70	200
macro avg	0.61	0.60	0.60	200
weighted avg	0.70	0.70	0.70	200

Nilai akurasi sesudah ditambahkan aturan baru :

	precision	recall	f1-score	support
ABBR	0.00	0.00	0.00	4
DESC	0.72	0.57	0.64	37
ENTY	0.52	0.67	0.59	51
HUM	0.78	0.76	0.77	50
LOC	0.85	0.88	0.86	32
NUM	0.83	0.77	0.80	26
accuracy			0.70	200
macro avg	0.62	0.61	0.61	200
weighted avg	0.71	0.70	0.70	200

Dapat dilihat perbedaan nilai dari keduanya, lebih dari 50% nilai dari yang atas naik setelah diberi aturan baru, nilai lainnya tidak mengalami penurunan hanya ada kestabilan nilai. Jadi, menambah aturan baru secara manual seperti di atas memiliki peran yang baik karena nilai akurasi presisi recall dan f1-score nya meningkat walau hanya sedikit.

4. Kira-kira fitur apa lagi yang dapat ditambahkan untuk meningkatkan kinerja identifikasi tipe question? Lengkapi jawaban poin ini dengan referensi.

Fitur yang dapat ditambahkan untuk meningkatkan kinerja identifikasi tipe question adalah fitur ekstraksi pertanyaan agar lebih jelas lagi pertanyaan tersebut masuk dalam kategori mana dan label mana, sehingga dapat mendapatkan jawaban yang sesuai.

Referensi

https://lms.telkomuniversity.ac.id/pluginfile.php/550593/course/section/79911/14_sistem_tanya_jawab.pdf