Perankingan Jawaban yang Terklasifikasi pada Komunitas Tanya-Jawab dengan*Term Frequency* dan *Similarity Measure Features*

Ranking the Classified Answer from Questioning Answering Community Using Term Frequency and Similarity Measure Features

Ali Ridho Fauzi Rahman[#1], Ir.Moch Arif Bijaksana,M.T[#2], Ade Romadhony,S.T.,M.T[#3]

#School of Computing, Telkom University
Jl. Telekomunikasi No. 01, Terusan Buah Batu, Bandung, Jawa Barat, Indonesia
[1]aliwumpa@gmail.com
[2]arifbijaksana@gmail.com
[3]aderomadhony@gmail.com

Abstrak

Banyak sekali orang bertukar informasi melewati *forum online*.Salah satu forum yang menyediakan lahan untuk bertukar informasi seputar Negara Qatar yaitu *Qatar Living Website Forum* memiliki banyak sekali orang yang bertanya maupun menjawab mengenai hal-hal yang ada di sekitar Negara Qatar,namun banyak sekali jawaban dari *responden* yang tidak berkaitan dengan hal yang ditanyakan.Pada tugas akhir ini dilakukan penelitian perankingan jawaban menggunakan metode *Term Frequency* dan *Similarity Measure Features.*Metode *Term Frequency* ini memiliki keunggulan untuk menghitung *score* kalimat jawaban yang akan dirangkingkan berdasarkan banyaknya jumlah term yang ada pada setiap kalimat jawabannya,sedangkan *Similarity Measure Features* dibagi menjadi dua fitur yaitu *Semantic Similarity* dan *Jaccard Similarity* memiliki keunggulan untuk menghitung besarnya kesimilaritasan antar kalimat berdasarkan kemiripan makna dan konten kalimat tersebut.Perankingkan jawaban dilakukan berdasarkan *score Term Frequency* nya dan tingkat keakurasian *Similarity Measure Features* nya dengan tahapan *Preprocessing*,*Feature Calculation*,dan *Ranking the Result with MAP evaluation*.Dari pengujian yang dilakukan,fitur yang memiliki tingkat kelayakan untuk merankingkan jawaban dengan MAP sebesar 80% adalah fitur *Semantic Similarity* yang merupakan salah satu fitur dari *Similarity Measure Features*.

Kata kunci : *Semantic Similarity,Jaccard Similarity,Term Frequency,Questioning Answering*

1.Pendahuluan

Ketersediaan perangkat beserta koneksi untuk mengakses jaringan internet pada zaman kini semakin meningkat.Dengan begitu kita dapat mencari informasi tanpa terikat waktu dan tempat .Banyak sekali orang mencari maupun bertukar informasi melalui forum-forum yang berada di internet,baik itu informasi untuk bisnis,tempat rekreasi,entertainment dan lain sebagainya.Salah satunya adalah Qatar Living Website Forum yang menyediakan informasi seputar tempat yang berada di Qatar.Setiap jawaban yang diberikan orang di forum tentunya berbeda-beda dan terdapat kemungkinan jawaban tersebut tidak mempunyai keterkaitan dengan pertanyaan yang diberikan.Oleh karena itu,dibutuhkan suatu system yang dapat membantu merangking istilah paling mendekati dengan pertanyaan yang diajukan di dalam Komunitas Tanya-Jawab.Salah satu metode yang dapat membantu memecahkan masalah ini adalah *Term Frequency* dan *Similarity Measure Features* dimana *Term Frequency* ini dapat menghitung *score* kalimat jawaban dari banyaknya term yang ada dari setiap jawaban yang diberikan oleh *responden* sedangkan *Similarity Measure Features* yang dibagi menjadi *Semantic Similarity* dan *Jaccard Similarity* dapat menghitung kesimilaritasan antar kalimat dari segi makna dan konten kalimat jawaban dengan pertanyaan tersebut.

Urutan proses yang dilakukan dalam penelitian ini adalah preprocessing data,pengklasifikasian jawaban,pembobotan jawaban dan yang terakhir adalah perankingan pada setiap jawaban.Pada pembobotan jawaban akan di terapkan *Term Frequency* dan *Similarity Measure Features*.Kegunaan dari *Term Frequency* ini untuk menghitung kemunculan term atau kata yang muncul pada suatu dokumen sehingga jumlah term jawaban telah dilihat

berdasarkan pertanyaan yang berkaitan,sedangkan *Similarity Measure Features* untuk menghitung seberapa besar kemiripan atau keterkaitan antara pertanyaan dan jawaban yang telah diberikan.

2.Studi Literatur

Perankingan jawaban pada dataset *Qatar Living Website Forum* pada penelitian kali ini akan menggunakan fitur *Term Frequency,Semantic Similarity,*dan *Jaccard Similarity* yang akan dijelaskan secara singkat di bagian Studi Literatur ini beserta perhitungan *Mean Average Precision*(MAP) nya.

Term Frequency

Proses perhitungan jumlah term yang ada dengan inisialisasi t di dalam dokumen d sehingga dilambangkan TF(t,d).Namun jumlah *term* yang sama dan muncul pada satu dokumen yang sama tidak akan mempengaruhi tingkat relevansi pada perhitungan *term frequency*[10].Berikut merupakan rumus perhitungan *log frequency* untuk pembobotan *term* t di dalam document d dari *term frequency* :

$$w_{t,d} = \begin{cases} 1 + \log_{10} tf_{t,d} & tf_{t,d} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Rumus tersebut menjelaskan bahwa perhitungan menggunakan log untuk menghindari perhitungan bobot nilai yang menghasilkan nilai 0.Sedangkan untuk penambahan hasil log dengan 1 untuk menghindari pembobotan tidak terbatas atau *infinity*.Untuk perhitungan *scoring* terakhir antara *document-query pairing* dengan jumlah term yang muncul di keduanya menggunakan rumus berikut :

$$score = \sum (1 + \log tf_{t,d}) \quad (2)$$

Semantic Similarity

Fitur atau metode *Semantic Similarity* ini digunakan untuk mengetahui tingkat akurasi kemiripan atau relevansi berdasarkan jarak antar kata.Misalkan kata 1 dan kata 2 dengan inisial Wi dan Wj.Dalam kasus ini Wi merupakan *ancestor* dari Wj,kedua kata ini akan dihitung jarak relevansi nya menggunakan *semantics library* atau yang disebut dengan *wordnet* lalu akan dibagi dengan nilai kedalaman maksimum *tree semantic library (depth of tree)* nya berikut merupakan rumus perhitungan *semantic similarity* tersebut[11] :

$$sim(c_1, c_2) = \begin{cases} 1 - dist(c_1, c_2)/D & c_1 \neq c_2 \\ 0 & c_1 = c_2 \end{cases} \quad (3)$$

Di dalam penelitian tugas akhir ini saya menggunakan fitur library dari WS4J dengan metode perhitungan *Wu and Palmer*.Metode Semantic Similarity *Wu and Palmer* ini melihat kesimilaritasan antar kata berdasarkan kedalaman LCS(*Lowest Common Subsumer*) dan jalur terpendek.Proses perhitungan yang dilakukan oleh *Wu and Palmer* ini adalah mencari jalur terpendek dari setiap kata,lalu jalur yang terbentuk dari kedua kata itu digabungkan untuk mencari *sense* yang sering muncul dari gabungan jalur tersebut.Rumus perhitungan untuk *Wu and Palmer* ini adalah

$$wup(c_1, c_2) = \frac{2 depth(lcs(c_1, c_2))}{\min(depth(c_1)) + \min(depth(c_2))} \quad (4)$$

Jaccard Similarity

Jaccard Similarity merupakan salah satu metode mengukur kemiripan atau relevansi antar 2 kalimat berdasarkan irisan kata.Hasil nilai pengukuran metode *Jaccard Similarity* berada di rentang nilai 0 dan 1.Bila hasil pengukuran mendekati angka 0 maka kedua kalimat itu memiliki nilai **ketidakmiripan** yang besar dan sebaliknya bila hasil pengukuran mendekati angka 1 maka kedua kalimat itu memiliki nilai **kemiripan** yang besar.Berikut dibawah ini merupakan rumus perhitungan *Jaccard Similarity*[12] :

$$\text{Jaccard}(S,T) = \frac{|S \cap T|}{|S \cup T|} \qquad (5)$$

Rumus *Jaccard Similarity* diatas menjelaskan bahwa irisan elemen-elemen kata yang ada di dokumen S dengan dokumen T dibagi dengan gabungan seluruh elemen-elemen kata yang ada di dokumen S dan dokumen T.

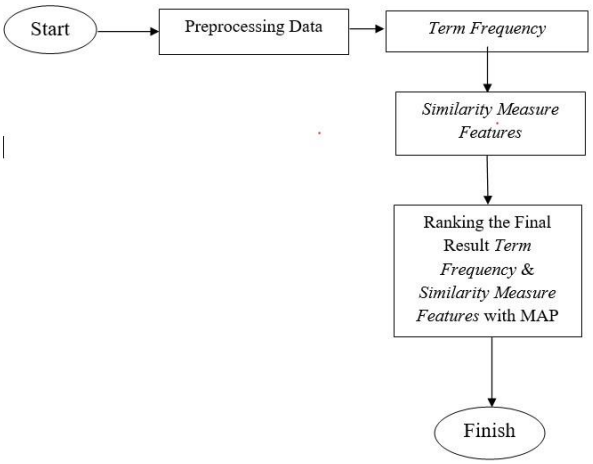
Mean Average Precision

Mean Average Precision atau yang disingkat dengan MAP adalah metode perhitungan untuk menghitung rata-rata dari *Average Precision*.Pengertian dari *Average Precision* itu sendiri adalah jumlah nilai *Precision* berdasarkan obyek terpilih yang bernilai *true/relevant* dibagi dengan jumlah semua item terpilih yang bernilai *true/relevant* seperti yang dijelaskan didalam teori *Information Retrieval*.Rumus perhitungan untuk MAP adalah sebagai berikut :

$$\text{MAP} = \frac{\sum \text{Average Precision}}{\text{Jumlah Average Precision}} \qquad (6)$$

3.Perancangan

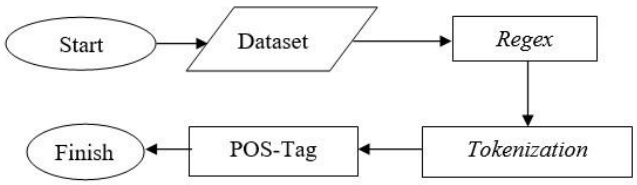
Perancangan sistem perankingan jawaban dilakukan melalui tahap *Preprocessing,Feature Calculation,*dan *Ranking the result based on MAP.*Berikut merupakan gambar alur sistem tersebut.



Gambar 1. *Flowchart* Sistem

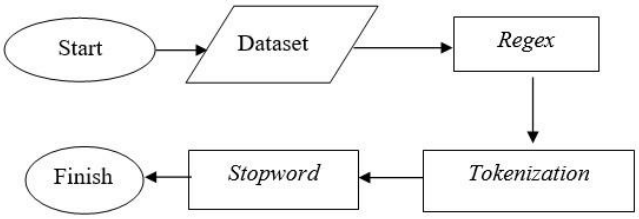
Preprocessing

Pada tahap ini sistem melakukan 2 preprocessing yang berbeda yakni sebagai berikut.



Gambar 2. *Flowchart Preprocessing* untuk *Semantic Similarity*

Proses perhitungan akurasi dengan *semantic similarity* ini membutuhkan POS-TAG karena penglabelan per kata untuk setiap kalimat berperan besar didalam metode perhitungan ini.Penglabelan per kata ini untuk melihat kata apa saja yang bertipe subjek,*verb*,dan objek yang nantinya akan dipakai untuk menghitung kesimilaritasan kata antar tipe label yang sama.



Gambar 3. *Flowchart Preprocessing* untuk *Term Frequency* dan *Jaccard Similarity*

Proses perhitungan metode *Term Frequency* dan *Jaccard Similarity* tidak perlu menggunakan POS-TAG karena kedua fitur ini melakukan proses perhitungan berdasarkan jumlah *term* dan irisan kata nya tanpa perlu melihat kata tersebut bertipe subjek,*verb*,ataupun objek.*Preprocessing* menggunakan *Stopword* ini untuk menghilangkan kata yang tidak baku dan tidak memiliki informasi apapun di dalam kata tersebut.Oleh karena itu *stopword* ini tidak digunakan di dalam *preprocessing* untuk *Semantic Similarity* karena dapat menghilangkan kata bertipe subjek ataupun objek di dalam kalimat jawaban yang akan di rankingkan.

Regex

Metode *preprocessing Regex* memiliki tujuan untuk menghilangkan simbol-simbol yang tidak diinginkan seperti @,^,%,# dan lain sebagainya dari kalimat jawaban yang akan di proses oleh sistem.

Contoh :

Kalimat : Try Both ;) me just trying to be helpful

Menjadi : Try Both me just trying to be helpful

Tokenization

Metode *preprocessing Tokenization* memiliki tujuan untuk memilah kalimat menjadi “token” per kata.

Contoh :

Kalimat : Try Both me just trying to be helpful

Menjadi : “Try”,”Both”,”me”,”just”,”trying”,”to”,”be”,helpful”

Part Of Speech-Tag

Metode preprocessing POS-Tag ini memiliki tujuan untuk memberikan label untuk setiap kata tunggal berdasarkan jenis katanya.Berikut merupakan daftar label yang ada di dalam POS-Tag ini :

Tabel 1. Daftar label Penn Treebank P.O.S Tags

Number	Tag	Description
1.	CC	Coordinating conjunction
2.	CD	Cardinal number
3.	DT	Determiner
4.	EX	Existential there
5.	FW	Foreign word
6.	IN	Preposition or subordinating conjunction
7.	JJ	Adjective
8.	JJR	Adjective, comparative
9.	JJS	Adjective, superlative
10.	LS	List item marker
11.	MD	Modal
12.	NN	Noun,singular or mass
13.	NNS	Noun,plural
14.	NNP	Proper noun,singular
15.	NNPS	Proper noun,plural
16.	PDT	Predeterminer
17.	POS	Possessive ending
18.	PRP	Personal pronoun
19.	PRP$	Possessive pronoun
20.	RB	Adverb
21.	RBR	Adverb,comparative
22.	RBS	Adverb,superlative
23.	RP	Particle
24.	SYM	Symbol
25.	TO	To
26.	UH	Interjection
27.	VB	Verb,base form
28.	VBD	Verb,past tense
29.	VBG	Verb,gerund or present participle
30.	VBN	Verb,past participle
31.	VBP	Verb,non-3rd person singular present
32.	VBZ	Verb,3rd person singular present
33.	WDT	Wh-determiner
34.	WP	Wh-pronoun
35.	WP$	Possessive wh-pronoun
36.	WRB	Wh-adverb

Contoh :

Kata : “Try”,”Me”

Menjadi : “Try_VB”,”Me_PRP”

Stopword Removal

Metode *preprocessing stopword removal* bertujuan menghilangkan kata imbuhan maupun kata-kata yang tidak deskriptif dan tidak memberikan informasi dari kalimat dan menyisakan kata-kata baku di dalam kalimat tersebut.

Contoh :

Kalimat : Try Both me just trying to helpful

Menjadi : helpful

Term Frequency

Metode perhitungan ini melihat terlebih dahulu jumlah *term* yang ada pada kalimat jawaban yang akan dihitung *score Term Frequency* nya.

Tabel 2. Daftar label Penn Treebank P.O.S Tags

term pertanyaan ke 1	*Frequency term* yang ada di jawaban ke 1
lot	1
massage	2
center	1
good	0
cost	0

Berdasarkan kasus diatas maka total term yang berada di kalimat jawaban ke 1 dilihat dari *distinct term* pertanyaan ke 1 adalah 4,lalu kita akan menghitung Bobot beserta Skor dari total term sebagai berikut :

$W(t,d) = 1+\log(4) = 1.60205$

Score $= 1+\log(W(t,d)) = 1+\log(1.60205) = 1.20467$

Setelah itu akan dilakukan perankingan dengan *sorting* beserta pengklasifikasian label berdasarkan *Scoring* fitur *Term Frequency*.

Threshold nilai untuk label “*Good*” adalah lebih dari sama dengan 3.0 (*result* >= 3.0),untuk label “*Potentially Useful*” adalah lebih dari 1.0(*result* > 1.0) dan selain dari itu adalah “*Bad*”.Pengklasifikasian label ini digunakan untuk perhitungan *precision* fitur ini.

Tabel 3.Hasil *sorting* dan pengklasifikasian label berdasarkan *score term frequency*

Kalimat Pertanyaan	Kalimat Jawaban	*Score* TF	Label
lot massage center massage center goodand cost	lot massage parlorall d city guessand d filipino massage center terramax gud d service excellentbut m curious thoughhavent facial center facial spa hereor call	4.05317806827809	Good
lot massage center massage center goodand cost	female places service women highly recommend biobil spa city center hour traditional aromatherapy massage oil 150 lady siam massage toys 435 4115 hour traditional thai massage 120	2.74127631137502	PotentiallyUseful
lot massage center massage center goodand cost	good facial centresspastry marriott ritz lovely spa expensive place opened opp city centre called bio somethingthey specialise stuffagain expensive lady siam centre toys alsadd area good toothey specialise thai	1.74127631137502	PotentiallyUseful
lot massage center massage center goodand cost	good thai massage place thai restaurant front doha clinic time driving closed filipino massage centre najma c d rings	1.52658903413904	PotentiallyUseful
lot massage center massage center goodand cost	gringer international massage centre alkinana street separate entrances men women kerala india good personally avail facility doctor centre consult complaints onehour full body massage qrs 100 telephone number	1.52658903413904	PotentiallyUseful
lot massage center massage center goodand cost	pls massage itthanks	1	Bad
lot massage center massage center goodand cost	message hmmmm recommend xxxxx semi guy shemales belive ahead decent al saad street al saad plaza coming traffic light bmw called kotakel good luck	0	Bad
lot massage center massage center goodand cost	guess wil lady shiam bio bilhhehehe hmmm doha	0	Bad
lot massage center massage center goodand cost	gringer belen bio bil shes amazing fun mc	0	Bad
lot massage center massage center goodand cost	crawl back rock troll	0	Bad

Jaccard Similarity

Untuk mendapatkan irisan digunakan metode pengecekan kata yang bersifat uniqe dari kalimat pertanyaan dengan *contains(word)* terhadap *Arrays.asList(answer)* sedangkan untuk nilai gabungan kalimat pertanyaan dengan jawaban digunakan jumlah *length* dari kalimat pertanyaan dan *length* dari kalimat jawaban.

$$\text{Result} = \frac{|A \cap B|}{A + B} \quad (7)$$

threshold nilainya sebagai berikut :

-*Result* >= 0.5 adalah "*Good*"

-0.5 > *Result* >= 0.3 adalah "*Potentially Useful*"

-*Result < 0.3* adalah "*Bad*"

Tabel 4.Hasil *sorting* dan pengklasifikasian label berdasarkan hasil *Jaccard Similarity*

Kalimat Pertanyaan	Kalimat Jawaban	Hasil JS	Label
hican place good massage drom philipinies yesterday massage biobil charged 300qr 01 hour bt totally waste pls advice philipinos	recommend good place head massages constantly migranes head massage medication job left drained	0.09375	Bad
hican place good massage drom philipinies yesterday massage biobil charged 300qr 01 hour bt totally waste pls advice philipinos	massages qatar waste money rub oil body deep tissue massage	0.0689655172413793	Bad
hican place good massage drom philipinies yesterday massage biobil charged 300qr 01 hour bt totally waste pls advice philipinos	masseuse good calling time home service philippines month vacation guess aromatherapy massage	0.0645161290322581	Bad
hican place good massage drom philipinies yesterday massage biobil charged 300qr 01 hour bt totally waste pls advice philipinos	call good 44410410	0.0454545454545455	Bad
hican place good massage drom philipinies yesterday massage biobil charged 300qr 01 hour bt totally waste pls advice philipinos	roy place contact number posted	0.0416666666666667	Bad
hican place good massage drom philipinies yesterday massage biobil charged 300qr 01 hour bt totally waste pls advice philipinos	massage doha money opinion merzam residencecream color building light blue glass window opposite center mega martbin mahmood price reasonable 150 hour therapist bali island working ritzcarlton hotel dubai	0.04	Bad
hican place good massage drom philipinies yesterday massage biobil charged 300qr 01 hour bt totally waste pls advice philipinos	magic touch abu hamour abu hamour petrol stnit cost 60qr hour lot qataris customers	0.0303030303030303	Bad
hican place good massage drom philipinies yesterday massage biobil charged 300qr 01 hour bt totally waste pls advice philipinos	call fast sooooo reservations	0	Bad
hican place good massage drom philipinies yesterday massage biobil charged 300qr 01 hour bt totally waste pls advice philipinos	roy beauty salon location	0	Bad

Semantic Similarity

Berikut merupakan tahapan perhitungan *Semantic Similarity* setiap kalimatnya :

-Mendapatkan kata dengan label bertipe Subjek,Verb,dan Objek dari kalimat pertanyaan dan jawaban untuk menghitung kesimilaritasan antar kata bertipe Subjek,Verb,dan Objek.Perlu diingat sistem memerlukan jenis-jenis label yang tergolong jenis Subjek,Verb,dan Objek.Berikut ini merupakan jenis penggolongan label yang saya buat :

-*Subject* = {"PRP","PRP$"}

-*Verb* = {“VB”,”VBD”,”VBG”,”VBN”,”VBP”,”VBZ”}

-*Object* = {“NN”,”NNS”,”NNP”,”NNPS”};

Tabel 5. Hasil penglabelan kata pertanyaan dan jawaban

Word Question	Label	Word Answer	Label
qatar	NN	you	PRP
philipinos	NNS	kahrama	NN
advise	VB	are	VBP
them	PRP	massages	NNS

-Kesimilaritasan antar kata bertipe Subjek,Verb,dan Objek ini memakai metode perhitungan *Wu and Palmer* sebagai berikut :

1.Perhitungan antar kata Subjek

$$\text{sim}(w_1, w_2) = \frac{2 \cdot \text{depth}(\text{lcs}(w_1, w_2))}{\min(\text{depth}(w_1)) + \min(\text{depth}(w_2))} \quad (8)$$

2.Perhitungan antar kata Verb

$$\text{sim}(w_1, w_2) = \frac{2 \cdot \text{depth}(\text{lcs}(w_1, w_2))}{\min(\text{depth}(w_1)) + \min(\text{depth}(w_2))} \quad (9)$$

3.Perhitungan antar kata Objek

$$\text{sim}(w_1, w_2) = \frac{2 \cdot \text{depth}(\text{lcs}(w_1, w_2))}{\min(\text{depth}(w_1)) + \min(\text{depth}(w_2))} \quad (10)$$

Tabel 6. Contoh hasil kesimilaritasan antar kata Subjek,Verb,dan Objek

sim(w1, w2)	Hasil	sim(w1, w2)	Hasil	sim(w1, w2)	Hasil
(me,you)	0.956719	(are,do)	0.913243	(qatar,philiphinos)	0.86039
(I,them)	0.765362	(visited,give)	0.657812	(job,visa)	0.42344
(they,it)	0.534231	(feed,talked)	0.329311	(circumstance,issued)	0.35294
(we,you)	0.520554	(put,care)	0.258938	(school,massages)	0.14345

-Setelah mendapatkan similaritas antar kata per label nya,sistem dapat menghitung kesimilaritasan antar kalimat dengan penjumlahan similaritas Sim(S1,S2),Sim(V1,V2),dan Sim(O1,O2) dibagi dengan 3[12].Berikut merupakan rumus perhitungannya :

$$\text{Sim}(\text{sentence}_1, \text{sentence}_2) = \frac{\text{Sim}(S_1, S_2) + \text{Sim}(V_1, V_2) + \text{Sim}(O_1, O_2)}{3} \quad (11)$$

Threshold nilai yang digunakan untuk pengklasifikasian label hasil perhitungan fitur ini adalah sebagai berikut :

- *Result* >= 0.5 adalah "*Good*"
- 0.5 > *Result* >= 0.3 adalah "*Potentially Useful*"
- *Result* < *0.3* adalah "*Bad*"

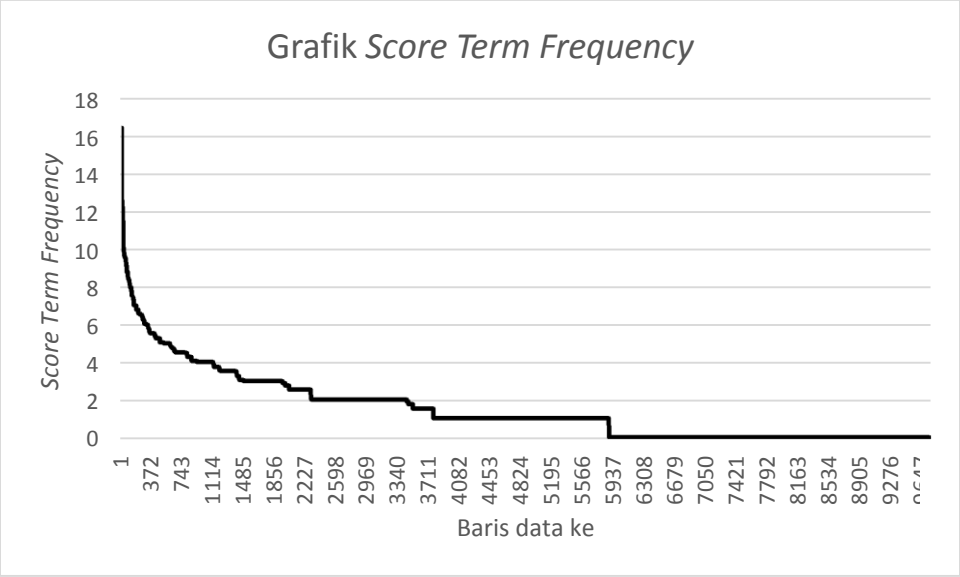
Tabel 7.Hasil kesimilaritasan kalimat pertanyaan dengan jawaban

Kalimat Pertanyaan	Kalimat Jawaban	Hasil WS	Label
dear ql members good morning moved dubai admit child grade cbse syllabi birla good doha modern indian school grateful members kids school give feed back admissions easily hope ur responses advance	telling mes indian school good andi xiith grade student	0.840123	Good
dear ql members good morning moved dubai admit child grade cbse syllabi birla good doha modern indian school grateful members kids school give feed back admissions easily hope ur responses advance	darude bashing i am just interested in getting child doha modern good high fee structure birla topic family joining weeks time make decision	0.732501	Good
dear ql members good morning moved dubai admit child grade cbse syllabi birla good doha modern indian school grateful members kids school give feed back admissions easily hope ur responses advance	mes worst birla is the best this is what people say in doha d	0.704567	Good
dear ql members good morning moved dubai admit child grade cbse syllabi birla good doha modern indian school grateful members kids school give feed back admissions easily hope ur responses advance	ajnas feedback contact school staff check admissions seat availability scenario	0.345342	PotentiallyUseful
dear ql members good morning moved dubai admit child grade cbse syllabi birla good doha modern indian school grateful members kids school give feed back admissions easily hope ur responses advance	compare marks activities students mes top gulf	0.310567	PotentiallyUseful
dear ql members good morning moved dubai admit child grade cbse syllabi birla good doha modern indian school grateful members kids school give feed back admissions easily hope ur responses advance	worry seats seats accomodate coz minimum 80 students class hehehe	0.172534	Bad
dear ql members good morning moved dubai admit child grade cbse syllabi birla good doha modern indian school grateful members kids school give feed back admissions easily hope ur responses advance	confused response acquaintances short stay giving extra activities coaching birla	0.165553	Bad
dear ql members good morning moved dubai admit child grade cbse syllabi birla good doha modern indian school grateful members kids school give feed back admissions easily hope ur responses advance	School bashing imgassistnid73057titledesclinknonealignleftwidthheight0	0.140322	Bad

dear ql members good morning moved dubai admit child grade cbse syllabi birla good doha modern indian school grateful members kids school give feed back admissions easily hope ur responses advance	max 42class	0	Bad
dear ql members good morning moved dubai admit child grade cbse syllabi birla good doha modern indian school grateful members kids school give feed back admissions easily hope ur responses advance	problem	0	Bad

Evaluasi

1.Berikut hasil *Term Frequency* berdasarkan skenario pengujian label "*Good*"



Gambar 4. Grafik hasil *Term Frequency*

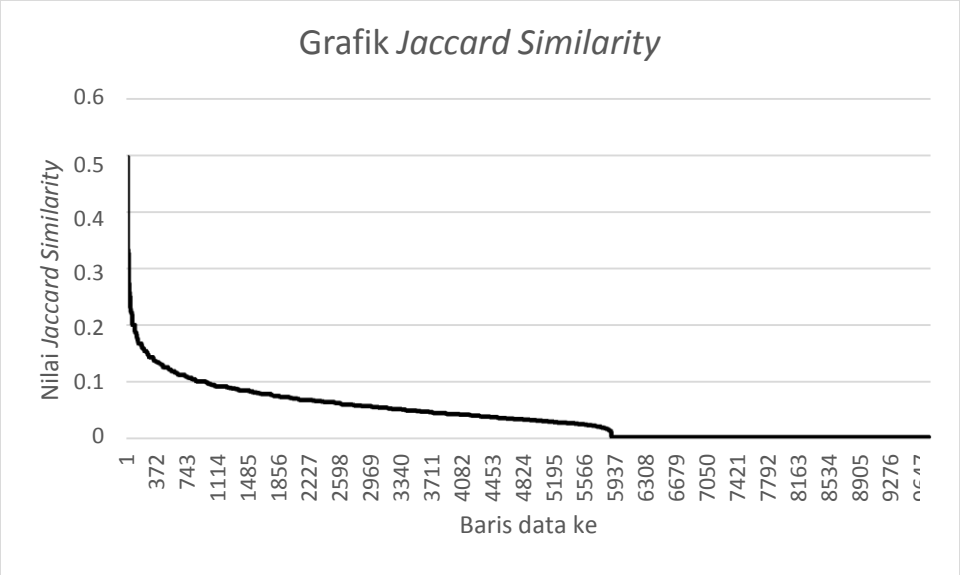
Dapat dilihat hasil perhitungan fitur yang bernilai "*Good*" (result >=3.0 hanya sampai *index* baris pertanyaan 1381 dan sekitar 8500 baris lainnya adalah *"Potentially Usefull"* dan *"Bad*".Bila dilihat dari penentuan nilai *true positive* dan *true negative* ,hanya label "*Good*" dalam fitur ini yang menjadi pembanding dengan label dataset bertipe *"Good"* dan "Bad" maka hasil dari *precision* untuk fitur ini adalah sebagai berikut :

Tabel 8. Hasil Perhitungan jumlah nilai *true positive* dan *false positive* dari 1381 baris berlabel "*Good*"

Jumlah baris bernilai TP	Jumlah baris bernilai FP	Jumlah baris lainnya
346	897	138

$$\text{[illegible]} = \frac{\text{[illegible]}}{\text{[illegible]}} = \frac{346}{346 + 897} = \frac{346}{1243} = 0{,}27835880$$

2.Berikut merupakan hasil *Jaccard Similarity* berdasarkan skenario pengujian label "*Good*"



Gambar 4. Grafik hasil *Jaccard Similarity*

Untuk setiap nilai *precision* yang dimiliki tiap dokumen *relevant* akan dijumlahkan seperti contoh hasil berikut ini :

Tabel 4-3 Contoh hasil Perhitungan *Jaccard Similarity* untuk jumlah nilai *precision* tiap dokumen *relevant*

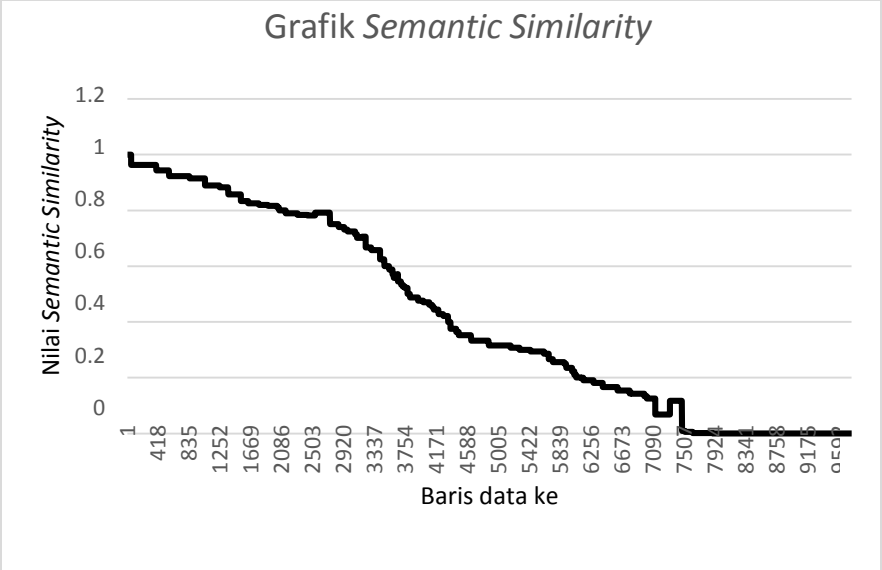
Hasil *Jaccard*	Label Fitur
0.5	*Good*

Setelah tiap nilai fitur *precision* dokumen *relevant* dijumlahkan lalu dibagi dengan jumlah data dokumen *relevant* nya untuk menghitung nilai *Mean Average Precision* global nya.Jumlah perhitungan nilai fitur dari 10 ribu baris yang berlabel "*Good*" adalah 0.5,sedangkan jumlah data dokumen berlabel "*Good"* adalah 1.Maka *Mean Average Precision* nya adalah

$$\text{Mean Average Precision} = \frac{\sum \text{Jaccard Similarity}}{\text{Jumlah Data}} = \frac{0.5}{1}$$

$$= 0{,}5$$

3.Berikut merupakan hasil *Semantic Similarity* berdasarkan skenario pengujian label "*Good*"



Gambar 4. Grafik hasil *Semantic Similarity*

Contoh tabel hasil perhitungan :

Tabel 4-4 Contoh hasil Perhitungan *Semantic Similarity* untuk jumlah nilai *precision* tiap dokumen *relevant*

Hasil *Semantic Similarity*	Label Fitur
1	*Good*
0.923077	*Good*
0.889231	*Good*
0.444444	*PotentiallyUseful*
Jumlah Dokumen *relevant* :	1+0.923077+0.889231 = 2.812308

Jumlah tiap perhitungan nilai dari fitur *Semantic Similarity* dari 10 ribu baris yang berlabel "*Good*" adalah 3153.487,sedangkan jumlah data dokumen berlabel "*Good"* adalah 3900.Maka perhitungan *Mean Precision Average* nya adalah

$$\text{[illegible]} = \frac{\sum \text{[illegible]}}{\text{[illegible]}} = \frac{3153.487}{3900}$$

$$= 0{,}808586$$

Dari hasil pengujian dengan perhitungan MAP berdasarkan data berlabel "*Good*" adalah fitur *Term Frequency* memiliki *precision* yang paling kecil dengan nilai MAP 0,27835880 dan urutan *precision* kedua adalah *Jaccard Similarity* dengan nilai MAP 0,5 lalu dengan *precision* terbesar adalah *Semantic Similarity* dengan nilai MAP 0,808586.

Kesimpulan

Berikut merupakan kesimpulan dari hasil pengujian tersebut :

a.)Perhitungan dengan *Semantic Similarity* antar kalimat memiliki akurasi paling tinggi dengan *Mean Precision Average* sebesar 80% bila dibandingkan fitur *Jaccard Similarity* dan *Term Frequency*,karena perhitungan ini memiliki perhitungan yang cukup kompleks dengan melibatkan tiap kesimilaritasan kata antar Subjek,Verb,dan Objek.

b.)Perhitungan dengan *Term Frequency* memiliki perhitungan *precision* paling rendah yaitu 27,8% dikarenakan jumlah data yang berlabel "*Good"* hanya 1/5 dari keseluruhan data yang ada.Ini disebabkan oleh kalimat jawaban yang variatif dan kebanyakan tidak menggunakan kembali *term* yang bersangkutan dengan kalimat pertanyaan yang di berikan oleh responder *Qatar Living Website Forum.*

c.)Perhitungan dengan *Jaccard Similarity* memiliki perhitungan *precision* 50 % namun jumlah data berlabel good hanyalah 1 dari 10 ribu baris data yan ada.Ini disebabkan penggunaan metode yang tidak cocok dengan bentuk dataset yang ada.

Saran

a.)Dapat menambahkan fitur similarity lain seperti *cosine similarity* dengan metode perhitungan yang hampir serupa namun berbeda dengan *jaccard similarity* dan mungkin akan berbeda hasil perhitungannya fiturnya.

b.) Melakukan perhitungan untuk Label "*PotentiallyUseful*" untuk melihat kecenderungan tingkat akurasi tersebut lebih mengarah ke Label "*Good*" atau "*Bad*" sehingga memungkinkan perhitungan sistem per fitur nya lebih akurat.

Daftar Pustaka

[1] Nakov, Preslav, et al. semEval-2015 Task 3: Answer Selection in Community Question Answering. Proceedings of the 9th International Workshop on Semantic Evaluation, SemEval. Vol. 15. 2015.

[2] Frequency Based Features Selection.http://nlp.stanford.edu/IRbook/html/htmledition/frequency-based-feature-selection-1.html.

[3] Information Retrieval. http://kholid.lecturer.pens.ac.id/KuliahLama/DSI/Day%206%20%20Information%20Retrieval.pdf.

[4] Evaluation in Information Retrieval. http://www-nlp.stanford.edu/IR-book/.

[5] Text Classification and Naive Bayes. http://www-nlp.stanford.edu/IR-book/.

[6] Collober,Ronan, et al. Journal of Machine Learning Research 12:Natural Language Processing(Almost) from Scratch. Proceedings of the NEC Laboratories America 2011.

[7] Tran, Quan Hung,et al.semEval-2015 Task 3 : Combining multiple features for Answer Selection in Community Question Answering

[8] Chaves, Marcirio Silveira, et al. Applying a Lexical Similarity Measure to Compare Portuguese Term Collections.Proceedings of the Ponitificia Universidade Catolica do Rio Grande do Sul – PUCRS

[9] Nicosia, Massimo,et al. semEval-2015 Task 3:Answer Selection for Community Question Answering – Experiment for Arabic and English

[10] Term-frequency-and-weighting. http://www-nlp.stanford.edu/IR-book/.

[11]Li Yuhua,McLean David,et al.Sentence Similarity Based on Semantic Nets and Corpus Statistics.IEEE Transactions On Knowledge and Data Engineering.Vol.18.2006.

[12]Liu Yuntong,Liang Yanjun.A Sentence Semantic Similarity Calculating Method Based On Segmented Semantic Comparsion.Journal of Theoretical and Applied Information Technology.Vol.48.2013.

[13]Maulana Akip,Bijaksana Arif,Mubarok Syahrul.Perancangan *Semantic Similarity based on Word Thesaurus* Menggunakan Pengukuran Omiotis untuk Pencarian Aplikasi pada I-GRACIAS.

[14]Voorhees Ellen.Doug Oard *Standford.edu.class Mean Average Precision.*2004