

POS Tagger Tweet Bahasa Indonesia

Yuda Munarko^{*1}, Yufis Azhar², Maulina Balqis³, Susi Ekawati⁴
^{1,2,3,4}Universitas Muhammadiyah Malang
yuda@umm.ac.id^{*}

Abstrak

Pada penelitian ini dilakukan investigasi POS Tagger dengan pendekatan Cyclic Dependency Network untuk data tweet dalam Bahasa Indonesia. Untuk koleksi tweet, digunakan tiga koleksi data, yakni tweet dengan gaya bahasa formal, informal dan gabungan. Sumber koleksi tweet formal adalah tweet dari akun berita, sedangkan koleksi tweet informal didapatkan dari akun umum. Adapun jenis tag yang digunakan berjumlah 41, dimana 35 adalah standar tag Bahasa Indonesia dan 6 adalah tambahan tag untuk twitter. Hasilnya adalah untuk koleksi data formal ketepatan deteksi mencapai 95,42%. Sedangkan untuk koleksi data informal dan gabungan ketepatannya mencapai 92,42% dan 90,69% secara berurutan. Kami juga mendapatkan hasil bahwa untuk tag yang sering muncul cenderung untuk memiliki nilai ketepatan yang tinggi juga, sedangkan tag yang kemunculannya lebih sedikit menyebabkan penurunan rata-rata ketepatan secara keseluruhan.

Kata kunci: Cyclic dependency network, POS Tagger

Abstract

In this research, we investigated POS Tagger with Cyclic Dependency Network approach for tweet data in Indonesian. The data consisted of three categories; tweets in formal language, tweets in informal language, and the mix of them. The formal tweets are obtained from news account, while the informal tweets are obtained from public account. There are 41 types of tag being used, in which 35 of them are the standard tag in Indonesian, and the other 6 types are additional tags for twitter. The finding reveals that in formal data, the detection accuracy is up to 95.42%. Meanwhile, in the informal and mixed data the detection accuracy are up to 92.42% and 90.69% respectively. We also found that the popular tag tend to have highest detection accuracy, while the less popular tag affects the decreasing of its detection accuracy.

Keywords: Cyclic dependency network, POS Tagger

1. Pendahuluan

Media sosial Twitter saat ini menjadi alternatif sumber informasi bagi masyarakat umum yang bersifat *realtime*. Contoh penggunaan praktisnya adalah untuk mendapatkan informasi lalu lintas, peristiwa penting, *event*, dan lain-lain. Sehubungan dengan hal tersebut, studi yang berhubungan dengan pengolahan data Twitter menjadi kegiatan yang sangat menarik. Selanjutnya, Twitter tidak hanya digunakan untuk mencari informasi saja, namun berkembang untuk beberapa kebutuhan, misalkan guna mengekstraksi pengetahuan dan juga meramalkan keadaan di masa yang akan datang

Maksud tujuan untuk menunjang kebutuhan tersebut, maka dikembangkan perangkat-perangkat *Natural Language Processing* (NLP), salah satunya adalah POS Tagger. POS Tagger merupakan proses klasifikasi kata suatu kalimat atau teks ke dalam *tag* tertentu. Studi awal dibidang ini salah satunya dilakukan oleh Marshall, 1987, menggunakan pendekatan probabilitas dan oleh Church, 1988, dengan model stokastik yang mempertimbangkan kaitan antar kata dari kanan ke kiri [1][2]. Kemudian ada juga karya Toutanova et al., 2003, menghasilkan salah satu pendekatan paling sukses dan banyak diadopsi hingga saat ini [3]. Perbedaan metode Toutanova dengan metode sebelumnya adalah diperhitungkannya kelas kata dan kata yang menghimpit kata yang akan dideteksi. Lebih jelas lagi, contoh ada rangkaian kata $w_0-w_1-w_2$ dimana tandanya adalah $t_0-t_1-t_2$, maka untuk menentukan t_1 , Toutanova mempertimbangkan w_0, t_0, w_1, w_2, t_2 , sedangkan dua metode sebelumnya hanya mempertimbangkan w_0, t_0 , atau w_2, t_2 saja.

Metode yang dikembangkan oleh Marshall, 1987, Church, 1988 dan Toutanova, 2003 dikembangkan dengan Korpus Bahasa Inggris. Dalam kasus POS Tagger Bahasa Indonesia, metode-metode tersebut akan memberikan hasil yang bagus, jika Korpus yang digunakan adalah Korpus Bahasa Indonesia. Hal inilah yang dilakukan oleh Pisceldo et al., 2009, dengan pendekatan probabilistik, Wicaksono & Purwariyanti, 2010, dan Hassan et al., 2011, yang mana kedua-duanya menggunakan Hidden Markov Model [4][5][6]. Untuk yang disebutkan pertama, metode yang digunakan pada dasarnya sama dengan Toutanova, 2000, menggunakan pendekatan maximum entropy [7]. Sedangkan yang disebutkan kedua dan ketiga memiliki kesamaan dengan Lee et al., 2000 [8]. Perbedaan keduanya adalah Hassan menggunakan data Mahasa Melayu, meskipun struktur kalimat Bahasa Melayu menyerupai dengan Bahasa Indonesia.

Penelitian POS Tagger khusus data Twitter mulai berkembang sekitar tahun 2011, misalnya yang dilakukan oleh Kevin et al., 2011, Foster et al., 2011, Derczynski et al., 2013, dan Zunchen et al., 2012 [9][10][11][12]. Meskipun fokus akhir dan pendekatan pada penelitian-penelitian berbeda-beda, namun salah satu pokok pembahasan utamanya adalah proses POS Tagger data Twitter yang tergolong rumit, karena tata bahasa dan penggunaan kata yang tidak formal. Pada penelitian Kevin et al., diformulasikan tambahan jenis *tag* khusus untuk Twitter sejumlah 6 *tag*. Hal yang sama dilakukan dengan menambahkan 6 *tag* tersebut pada *tweet* Bahasa Indonesia yang didefinisikan oleh Wicaksono & Purwariyanti, 2010 [5].

Karena penelitian-penelitian POS Tagger data Twitter sebagian besar berbasis Bahasa Inggris, maka perlu kiranya untuk melakukan penelitian yang sama tetapi dalam Bahasa Indonesia. Inilah yang menjadi kontribusi pertama pada penelitian ini. Kontribusi lainnya adalah tersedianya Korpus dan model untuk *tweet* Bahasa Indonesia. Disamping itu Korpus yang tersedia telah dibedakan menjadi dua kelompok, yakni untuk bahasa formal dan bahasa informal.

2. Metode Penelitian

2.1. Data dan Sumber Data

Data yang digunakan diambil dari situs Twitter dari dua jenis akun, seperti akun berita *online* dan akun umum. Pemilihan dua jenis akun ini dimaksudkan agar didapatkan data dengan jenis data bahasa formal dan informal. Dari setiap jenis data selanjutnya dibagi menjadi dua, yakni data untuk kepentingan pembuatan model, yaitu data latih dan data untuk menguji performa, disebut data uji.

Tabel 1. Data yang digunakan

jenis data	jumlah tweet		jumlah token	
	latih	uji	latih	uji
formal	4001	1000	59523	15998
informal	2000	500	32863	8771
gabungan	6001	1500	92386	24769

Tabel 1, menunjukkan jumlah data keseluruhan dan jumlah data uji maupun data latih yang digunakan. Selain itu data formal dan data informal selanjutnya digabung menjadi data gabungan. Berikut adalah contoh data twitter yang digunakan. Pertama adalah *tweet* formal dan selanjutnya adalah *tweet* informal.

@detikcom : Kebakaran Rumah di Pondok Bambu , 16 Unit Damkar Diterjankan
<http://detik.id/VjAuOQ>

@bentangpustaka Wihhh " @AuthorZACK : @bentangpustaka tips #Semangatpagi ala
 Aku adlh membaca novel kak Andrea Hirata :-) "

2.2. Proses Anotasi

Anotasi secara manual dilakukan oleh 3 orang ahli Bahasa, dimana 2 orang berfungsi sebagai anotator dan sisanya adalah validator. Mekanismenya, masing-masing anotator akan melakukan anotasi terhadap data yang sama, selanjutnya hasilnya dibandingkan antara satu dengan yang lain. Jika suatu kata memiliki *tag* yang sama, maka hasil tersebut dianggap valid. Jika hasil tidak sama, maka tugas validator untuk menentukan jenis *tag*. Adapun penggunaan

jenis *tag* pada penelitian ini, seperti yang didefinisikan oleh Wicaksono & Purwariyanti, 2010, dan Kevin et al., 2011. Sebagaimana ditampilkan Tabel 2. dan Tabel 3 [5][9].

Tabel 2. Jenis Label untuk Anotasi Bahasa Indonesia [5]

No	POS	Nama POS	Contoh	No	POS	Nama POS	Contoh
1	OP	Kurung buka	{[19	MD	Modal	Bisa
2	CP	Kurung tutup)]	20	CC	Kt hubung konjungtif	Dan, Atau, Tetapi
3	GM	Slash	/	21	SC	Kt hubung subordinatif	Jika, Ketika
4	;	Semicolon	;	22	DT	Penentu	Para, Ini, Itu
5	:	Colon	:	23	UH	Kata seru	Wah, Aduh, Oi
6	"	Kutipan	"	24	CDO	Angka ordinal	Pertama, Kedua
7	.	Penutup kalimat	!?	25	CDC	Angka kolektif	Bertiga
8	,	Koma	,	26	CDP	Angka utama	Satu, Dua
9	-	Dash	-	27	CDI	Angka tidak teratur	Beberapa
10	...	Ellipsis	...	28	PRP	Kata ganti orang	Saya, Kamu
11	JJ	Kata sifat	Kaya, Manis	29	WP	Kata tanya	Apa, Siapa
12	RB	Kata keterangan	Sementara, Nanti	30	PRN	Kata ganti angka	Kedua-duanya
13	NN	Kata benda umum	Mobil	31	PRL	Kata ganti lokasi	Sini, Situ, Sana
14	NNP	Kata benda spesifik	Bekasi, Indonesia	32	NEG	Negasi	Bukan, Tidak
15	NNG	Kata benda milik	Bukunya	33	SYM	Simbol	@#\$%^&
16	VBI	Kata kerja intransitif	Pergi	34	RP	Partikel	Pun, Kah
17	VBT	Kata kerja transitif	Membeli	35	FW	Kata asing	Foreign, Word
18	IN	Preposisi	Di, Ke, Dari				

Tabel 3. Jenis Label Tambahan untuk Anotasi Tweet [9]

No	POS	POS Name	Example
1	#	Hashtag, mengindikasikan topik atau kategori tweet	#sukaUMM
2	@	at-mention, mengindikasikan autor lain di tweet	@Muhammadiyah
3	~	Discourse marker, mengindikasikan kelanjutan satu atau beberapa tweet	RT dan : pada konstruksi retweet RT @detik :
4	U	URL atau alamat email	http://dlvr.it/B6R36z
5	E	Emoticon	:) :-) :b <3o__O
6	\$	numerik	123 tiga 12:30 XI

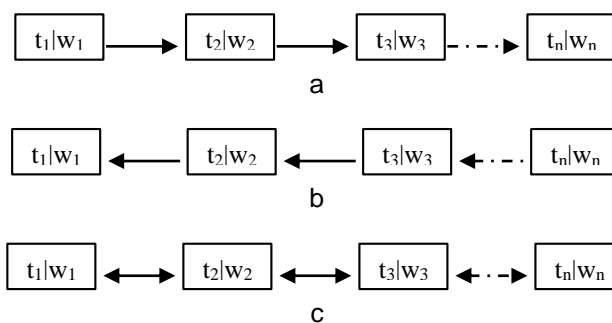
Contoh hasil dari proses anotasi tersebut ditampilkan pada Tabel 4, dimana dua yang pertama adalah data formal dan dua berikutnya adalah data informal.

Tabel 4. Contoh Tweet yang sudah Ditandai oleh Anotator

No	Tweet	Hasil Anotasi
1	@Motivasillmu : Jangan bersedih atas kesalahan . Karena kesalahan membuat kamu belajar bagaimana untuk menjadi orang yang benar .	@Motivasillmu/@ :/: Jangan/RB bersedih/VBT atas/NN kesalahan/NN ./ . Karena/CC kesalahan/NN membuat/VBT kamu/PRP belajar/VBI bagaimana/WP untuk/CC menjadi/VBT orang/NN yang/SC benar/JJ ./ .
2	@VIVAcoid : Kurangi Takaran Beras Miskin , Dituntut 7 Tahun Penjara http://dlvr.it/B6R36z	@VIVAcoid/@ :/: Kurangi/VBT Takaran/NN Beras/NN Miskin/JJ ./ , Dituntut/VBT 7/\$ Tahun/NN Penjara/NN http://dlvr.it/B6R36z/U
3	@bentangpustaka Wihhh " @AuthorZACK : @bentangpustaka tips #Semangatpag ala Aku adlh membaca novel kak Andrea Hirata :-)"	@bentangpustaka_@ Wihhh_UH " " @AuthorZACK_@ : : @bentangpustaka_@ tips_NN #Semangatpagi_# ala_NN Aku_PRP adlh_SC membaca_VBT novel_NN kak_PRP Andrea_NNP Hirata_NNP :-)_E "_ "
4	@dickymprasetyo : RT @Ayatie93 : gws Yasha ? ? "@fanadickyPKU Yasha harus sembuh " @GGSLoverssOFC : Yasha sekarat	@dickymprasetyo_@ : : RT_~ @Ayatie93_@ : : gws_FW Yasha_NNP ?_ ?_ . "@fanadickyPKU_@ Yasha_NNP harus_RB sembuh_VBI "_ " @GGSLoverssOFC_@ : : Yasha_NNP sekarat_JJ

2.2. Pendekatan *Cyclic Dependency Network*

Pendekatan menggunakan *cyclic dependency network* dengan mempertimbangkan multifitur dalam menentukan *tag*, merupakan salah satu yang terbaik dalam bidang ini. Toutanova et al., 2003 membuktikannya dengan melakukan percobaan terhadap data Penn Treebank Wall Street Journal dimana ketepatan deteksinya mencapai 97,24% [3]. Jika sebagian besar POS Tagger dalam menentukan *tag* hanya dipengaruhi oleh kata serta *tag* di sebelah kiri atau kanannya saja, maka pada pendekatan *cyclic dependency network* penentuan *tag* dipengaruhi oleh kata serta dua *tag* di kiri dan kanan. Supaya lebih jelas, Gambar 1 (a) menunjukkan bahwa t_2 dipengaruhi oleh t_1 , dan w_2 , dengan memanfaatkan metode Conditional Markov Model. Demikian juga pada Gambar 1 (b), t_2 dipengaruhi oleh t_3 , dan w_2 , dengan metode yang sama. Sedangkan pada *cyclic dependency network*, t_2 dipengaruhi oleh t_1 , t_3 , dan w_2 , dimana konsep ini disebut juga *Bidirectional Dependency Network* (BDN) yang didasari oleh pemahaman bahwa dalam identifikasi *tag* semua fitur yang berkorelasi langsung dengan suatu kata bisa dimanfaatkan dalam perhitungan.



Gambar 1. Dependency Network (a) CMM dari kiri ke kanan, (b) CMM dari kanan ke kiri, (c) Bidirectional Dependency Network

Permasalahan yang muncul dengan BDN adalah munculnya *cyclic dependency*, dimana akan sulit untuk menghitung, misalnya t_2 . Karena t_2 tergantung oleh nilai t_3 dan sebaliknya t_3 tergantung juga oleh nilai t_2 , dikarenakan nilai $P(t_2, t_3)$ ataupun $P(t_3, t_2)$ tidak mudah ditentukan. Untuk mengatasi masalah ini, berdasarkan Toutanova et al., 2003, maka digunakan Gibbs Sampling dan Local Maxent Model untuk memperkirakan probabilitas lokal tersebut [3].

3. Hasil Penelitian dan Pembahasan

Proses eksperimen diawali dengan membuat tiga jenis model dari tiga jenis data latih yang tersedia. Selanjutnya dilakukan proses *tagging* otomatis terhadap tiga jenis data uji, menggunakan tiga jenis model yang sudah dibuat. Hasil ketepatan deteksi tersebut disajikan di Tabel 5.

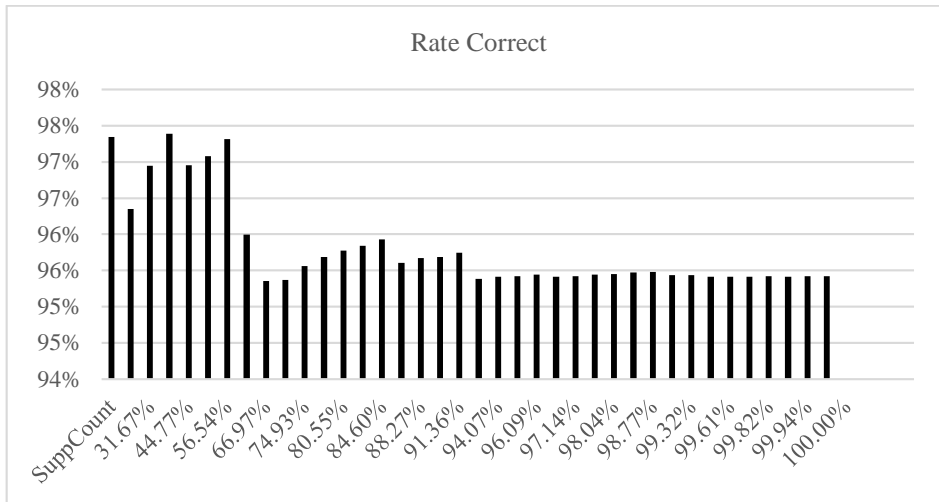
Tabel 5. Hasil POS Tagger

Data	Model	Tag	Kalimat	Kata Asing
formal	formal	95,42%	52,40%	71,28%
	informal	71,10%	1,60%	57,78%
	gabungan	93,63%	41,00%	71,73%
informal	formal	64.80%	0.20%	44.74%
	informal	92.42%	31.80%	73.55%
	gabungan	85.30%	9.80%	71.22%
gabungan	formal	84.58%	35.00%	55.48%
	informal	78.64%	11.67%	61.75%
	gabungan	90.69%	30.60%	71.48%

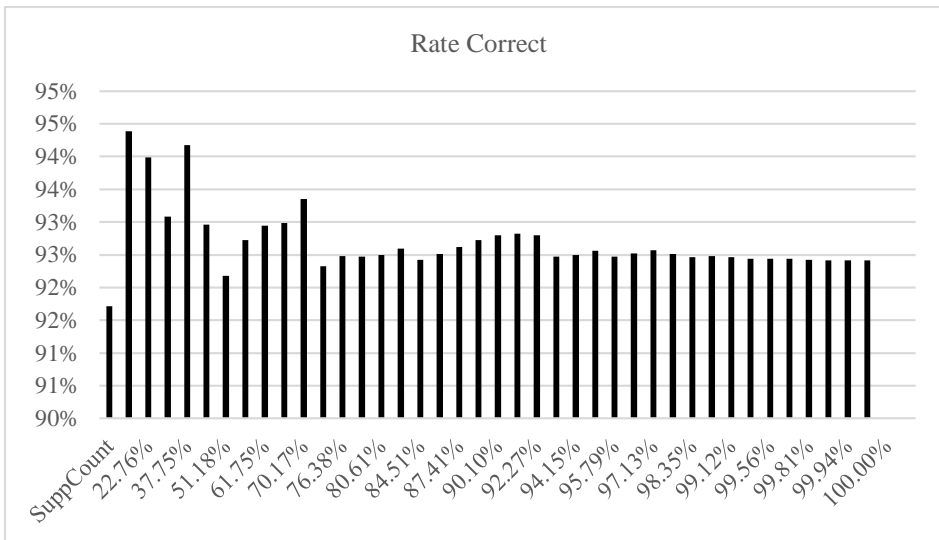
Tabel 6. Distribusi Kemunculan Tag

Formal			Informal			Gabungan		
Tag	Support	SuppCount	Tag	Support	SuppCount	Tag	Support	SuppCount
NN	20.27%	20.27%	NN	15.00%	15.00%	NN	18.40%	18.40%
VBT	11.41%	31.67%	@	7.75%	22.76%	VBT	9.93%	28.33%
@	6.59%	38.26%	RB	7.75%	30.51%	@	7.00%	35.33%
:	6.51%	44.77%	VBT	7.24%	37.75%	JJ	6.19%	41.52%
JJ	6.18%	50.95%	.	7.24%	44.99%	.	6.08%	47.60%
SC	5.59%	56.54%	JJ	6.19%	51.18%	SC	5.16%	52.76%
.	5.44%	61.99%	VBI	5.69%	56.87%	RB	4.98%	57.74%
NNP	4.98%	66.97%	IN	4.88%	61.75%	VBI	4.92%	62.66%
VBI	4.49%	71.47%	SC	4.37%	66.12%	:	4.68%	67.33%
RB	3.46%	74.93%	PRP	4.06%	70.17%	NNP	4.03%	71.36%
,	3.15%	78.08%	,	3.91%	74.09%	,	3.42%	74.78%
IN	2.47%	80.55%	NNP	2.29%	76.38%	IN	3.32%	78.11%
PRP	2.04%	82.59%	DT	2.14%	78.52%	PRP	2.75%	80.86%
CC	2.02%	84.60%	CC	2.09%	80.61%	CC	2.04%	82.90%

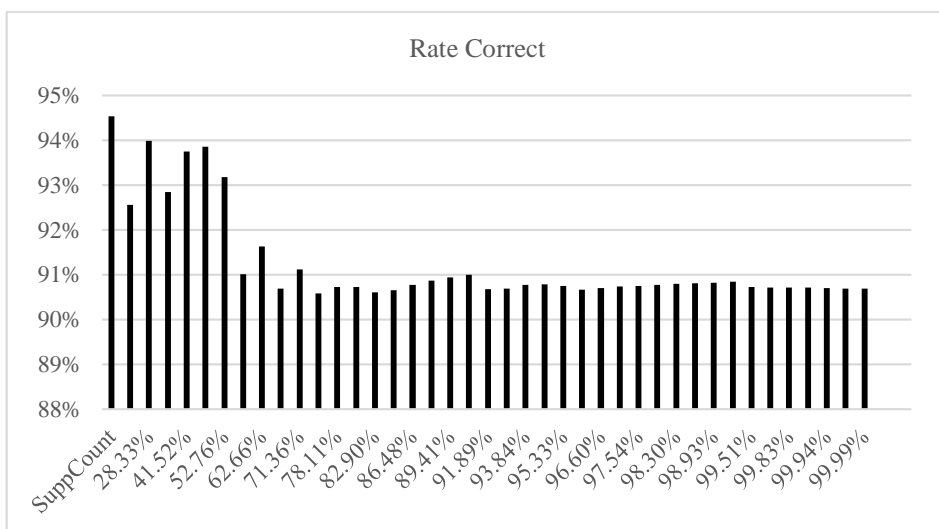
Tentu saja pendeteksian data uji formal menggunakan model formal, sedangkan data uji informal dideteksi menggunakan model informal, dan data uji gabungan menggunakan model gabungan memunculkan hasil tertinggi, yakni 95,42%, 92,42%, dan 90,69% pada deteksi *tag* disajikan secara berurutan. Hal ini adalah hasil yang wajar, karena POS Tagger bersifat domain spesifik, semakin spesifik kita mendefinisikan suatu model pada suatu domain maka hasil deteksi untuk domain tersebut akan semakin tepat. Masalah di Twitter adalah adanya pola penggunaan gaya bahasa yang beraneka ragam dan terkadang untuk deteksi *tweet* secara umum diperlukan model yang juga umum. Oleh sebab itu, pada penelitian ini diuji coba pula model gabungan. Model gabungan sudah memberikan hasil yang cukup bagus, meskipun pada ketepatan deteksi kalimat, performa model gabungan untuk mendeteksi data uji gabungan masih kalah dibandingkan model formal dalam mendeteksi data uji gabungan. Tinjauan pada segi deteksi kalimat menunjukkan data formal hasilnya mendekati hasil penelitian Toutanova et al., yakni kisaran 50% [3]. Memang untuk data informal dan gabungan persentasenya masih sekitar 30%, sehingga masih diperlukan peningkatan lebih lanjut. Dari apa yang sudah didapat, untuk selanjutnya penerapan teknik kluster besar kemungkinan dapat meningkatkan hasil deteksi. Implementasinya, pertama melakukan kluster terhadap kelompok data latih, lalu membangun model untuk setiap kluster, selanjutnya setiap *tweet* yang akan dideteksi ditentukan terlebih dahulu klusternya kemudian dideteksi menggunakan model yang sesuai.



a



b



c

Gambar 2. Pengaruh Tag yang Jarang Muncul (a) Formal, (b) Informal, (c) Gabungan

Tabel 6 menunjukkan pola persebaran jenis *tag*, dimana 80% lebih dari total keseluruhan kata didominasi oleh 15 jenis *tag* saja. Misalnya untuk *tag* NN di semua jenis data muncul dengan frekuensi tertinggi, kemudian disusul oleh VBT dan @. Sedangkan jenis *tag* lainnya juga hampir sama untuk setiap jenis data, namun dengan peringkat frekuensi yang berbeda-beda. Kondisi ini adalah sesuatu yang menarik, misalkan dimanfaatkan untuk melakukan klusterisasi data dan mendapatkan model yang berbeda-beda.

Gambar 2 menjelaskan rata-rata akurasi pendeteksian yang diukur dengan memperhitungkan jenis *tag* dari yang paling sering muncul sampai paling jarang muncul. Sebagai ilustrasi pada data formal, Gambar 1 (a), *bar* yang paling kiri adalah NN dengan nilai *support* 20,27%, sehingga *support* total adalah 20,27%. Kemudian *bar* kanannya adalah VBT dengan nilai *support* 11,41%, sehingga *support* total adalah 31,67%, demikian seterusnya. Dari Gambar 2 (a), Gambar 2 (b), dan Gambar 2 (c) tersebut terlihat bahwa semakin tinggi nilai *support* total nilai ketepatan deteksi akan semakin rendah. Kondisi ini mengindikasikan bahwa *tag* yang jarang muncul menyebabkan nilai akurasi deteksi menurun. Masalah tersebut kemungkinan disebabkan oleh jumlah data latih yang kurang, sehingga jumlah sample untuk *tag* yang jarang muncul juga sedikit. Untuk selanjutnya, sebaiknya jika jumlah data latih ditingkatkan lagi.

4. Kesimpulan

Berdasarkan pembahasan yang dilakukan, diketahui bahwa metode Cyclic Dependency Network memiliki tingkat akurasi cukup tinggi, yakni 95,42% data formal, 92,42% data informal, dan 90,69% data gabungan. Nilai tersebut sudah cukup baik, mengingat data Twitter terdiri dari kumpulan kata tidak baku, sehingga proses POS Tagging menjadi sulit. Ada indikasi bahwa nilai *support* yang besar akan berpengaruh terhadap nilai *confidence*. Sebenarnya ini berhubungan dengan jumlah data, semakin besar data uji kemungkinan dapat menaikkan nilai akurasi secara keseluruhan. Selain itu untuk pembuatan model ada baiknya data yang digunakan lebih dari 100.000, supaya bisa menampung banyak Corpora dan memberi gambaran yang lebih baik terhadap performa algoritma. Lebih lanjut lagi, data *training* yang besar mampu menciptakan model yang tangguh serta dapat digunakan oleh aplikasi lainnya. Berikutnya yang akan dilakukan adalah pembuatan POS Tagger multimodel menggunakan konsep klusterisasi dari data latih yang jumlahnya besar, sehingga *tweet* bisa dideteksi menggunakan model yang sesuai.

Referensi

- [1] Marshall I. "Tag selection using probabilistic methods". In Roger Garside, Geoffrey Sampson, and Geoffrey Leech, editors, *The Computational analysis of English: a corpus-based approach*. Longman, London. (1987): 42–65.
- [2] Church. KW. "A stochastic parts program and noun phrase parser for unrestricted text". In *ANLP 2*. (1988): 136-143.
- [3] Toutanova K., et al. "Feature-rich part-of-speech tagging with a cyclic dependency network." *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1*. Association for Computational Linguistics, 2003.
- [4] Pisceldo F., et al. "Probabilistic Part-of- Speech Tagging for Bahasa Indonesia". In *The Third International MALINDO Workshop, Colocated Event ACL-IJCNLP*. 2009.
- [5] Wicaksono AF., Purwariyanti A. "HMM based part-of-speech tagger for Bahasa Indonesia." *Fourth International MALINDO Workshop*, Jakarta. 2010.
- [6] Hassan M., et al. "Statistical malay part-of-speech (POS) tagger using Hidden Markov approach." *Semantic Technology and Information Retrieval (STAIR), 2011 International Conference on*. IEEE. 2011.
- [7] Toutanova K., Manning C. "Enriching the knowledge sources used in a maximum entropy part-of- speech tagger". In *EMNLP/VLC 1999*. (2000): 63–71.
- [8] Sang-Zoo Lee, Jun ichi Tsujii, and Hae-Chang Rim. Part- of-speech tagging based on Hidden Markov Model assuming joint independence. In *ACL 38 (2000)*; 263–169.
- [9] Kevin, et al. "Part-of-speech tagging for twitter: Annotation, features, and experiments." *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*. Association for Computational Linguistics. 2011.

- [10] Foster J., et al. "# hardtoparse: POS Tagging and Parsing the Twittersverse." proceedings of the Workshop On Analyzing Microtext (AAAI 2011). 2011.
- [11] Derczynski L., et al. "Twitter Part-of- Speech Tagging for All: Overcoming Sparse and Noisy Data". In Proceedings of the International Conference on Recent Advances in Natural Language Processing, ACL. 2013.
- [12] Zhunchen L., et al. "Improving Twitter Retrieval by Exploiting Structural Information." AAAI. 2012.