

Meningkatkan Peran Model Bahasa dalam Mesin Penerjemah Statistik (Studi Kasus Bahasa Indonesia-Dayak Kanayatn)

Herry Sujaini

Program Studi Teknik Informatika

Universitas Tanjungpura

Indonesia

hs@untan.ac.id

Abstrak-Sistem terjemahan mesin berbasis statistik menggunakan kombinasi satu atau lebih model terjemahan dan model bahasa. Meskipun ada banyak penelitian yang membahas peningkatan model terjemahan, masalah mengoptimalkan model bahasa untuk tugas penerjemahan tertentu belum banyak mendapat perhatian. Biasanya, model trigram digunakan sebagai model bahasa standar dalam sistem terjemahan mesin statistik. Dalam tulisan ini kami menerapkan 4 strategi eksperimen untuk melihat peran model bahasa yang digunakan dalam mesin terjemahan Indonesia-Dayak Kanayatn dan menunjukkan perbaikan pada sistem baseline dengan model bahasa standar.

Kata Kunci: terjemahan mesin, model bahasa, Indonesia-Dayak Kanayatn

1. Pendahuluan

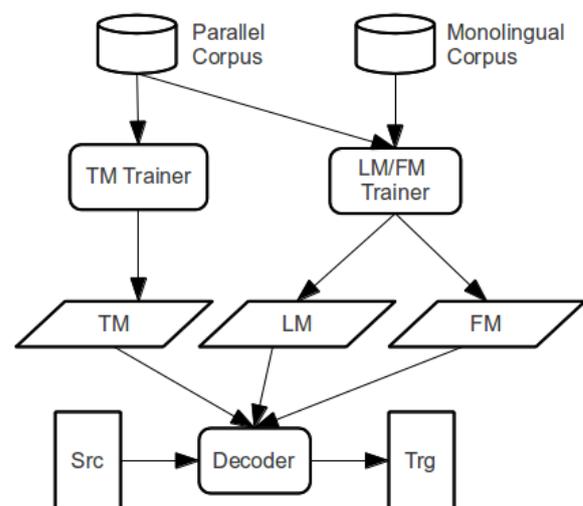
Mesin terjemahan mesin berbasis statistik atau dikenal dengan mesin penerjemah statistik (MPS) adalah sebuah paradigma penerjemahan mesin di mana terjemahan dihasilkan berdasarkan model statistik yang parameternya berasal dari analisis korpus bilingual (korpus paralel). Korpus didefinisikan sebagai koleksi atau sekumpulan contoh teks tulis atau lisan dalam bentuk data yang dapat dibaca dengan menggunakan seperangkat mesin dan dapat diberi catatan berupa berbagai bentuk informasi linguistik [1]. Lebih lanjut, McEnery merangkum kriteria korpus, yang telah menjadi kesepakatan banyak ahli, yakni: (1) dapat dibaca dengan menggunakan seperangkat mesin, (2) berupa teks otentik, (3) digunakan sebagai sampel, dan (4) mewakili bahasa atau variasi bahasa tertentu.

Tujuannya analisis korpus bilingual adalah untuk menemukan hipotesis terbaik \bar{e}^* setiap kalimat masukan f sebagai :

$$\bar{e}^* = \underset{\bar{e}}{\operatorname{argmax}} P(\bar{e}|f) = \underset{\bar{e}}{\operatorname{argmax}} P(f|\bar{e})P(\bar{e}) \quad (1)$$

$P(\bar{e}|f)$ adalah model terjemahan yang mengekspresikan kemungkinan probabilistik pada hubungan antara bahasa sumber dan bahasa target. $P(\bar{e})$ adalah model bahasa yang menentukan probabilitas *string* pada bahasa target yang biasanya menggunakan kata standar model trigram dari :

$$P(e_1, \dots, e_l) \approx \prod_{i=3}^l P(e_i|e_{i-1}, e_{i-2}) \quad (2)$$



Gambar 1. Arsitektur Mesin Penerjemah Statistik [3]

Di mana $\bar{e} = e_1, \dots, e_l$. Pada bentuk model trigram, setiap kata diprediksi berdasarkan riwayat dua kata sebelumnya.

Sebagian besar pekerjaan di MPS berkonsentrasi pada pengembangan model terjemahan yang lebih baik, *decoding algorithms*, atau *minimum error rate training* untuk MPS. Sedikit usaha telah dilakukan untuk memaksimalkan peran pemodelan bahasa untuk terjemahan mesin. Di bidang lain, terutama dalam pengenalan ucapan otomatis (ASR), ada banyak pekerjaan dalam pemodelan bahasa statistik yang membahasnya misalnya penggunaan kelas kata, adaptasi model bahasa, atau teknik estimasi probabilitas alternatif [2]. Tujuan dari penelitian ini adalah untuk meningkatkan

peran pemodelan bahasa yang pada gilirannya akan dapat meningkatkan akurasi dari hasil terjemahan suatu MPS.

Arsitektur mesin penerjemah statistik secara umum seperti terlihat pada Gambar 1. Sumber data utama yang dipergunakan adalah korpus paralel (*parallel corpus*) dan korpus monolingual (*monolingual corpus*). Proses *training* terhadap *parallel corpus* menghasilkan translation model (TM). Proses *training* terhadap bahasa target pada korpus paralel ditambah dengan korpus monolingual bahasa target menghasilkan model bahasa (*language model*) (LM), sedangkan fitur model (FM) dihasilkan dari bahasa target pada *parallel corpus* yang setiap katanya sudah ditandai dengan fitur linguistik seperti PoS, lemma, gender, proses pembentukan kata (morfem) dan lain-lain. TM, LM dan FM hasil proses di atas digunakan untuk menghasilkan *decoder*. Selanjutnya *decoder* digunakan sebagai mesin penerjemah untuk menghasilkan bahasa target dari *input* kalimat dalam bahasa sumber [3].

Jika dilihat dari arsitektur MPS, terlihat jelas bahwa data utama yang digunakan untuk menghasilkan model-model pada MPS adalah *parallel corpus*. *Monolingual corpus* dapat diperoleh dari *parallel corpus* pada sisi bahasa target walaupun biasanya diperbanyak lagi dari sumber-sumber lainnya.

Penelitian yang dilakukan untuk meningkatkan akurasi mesin penerjemah bahasa daerah tidak hanya pada area model bahasa. Mandira dkk. [4] melakukan perbaikan probabilitas *lexical* model untuk meningkatkan akurasi MPS bahasa Indonesia-Jawa sebesar 10,69%, Jarob dkk. [5] melakukan uji akurasi MPS bahasa Indonesia-Dayak Taman dengan penandaan kata dasar dan imbuhan yang menghasilkan peningkatan akurasi sebesar 0,36%, Nugroho dkk. [6] melakukan penelitian terhadap MPS bahasa Indonesia-Jawa menunjukkan hasil evaluasi 44,02 % untuk Indonesia Jawa dan 48,77 % untuk Jawa-Indonesia.

Beberapa penelitian untuk meningkatkan peran model bahasa telah dilakukan dalam berbagai bahasa dan berbagai metode, di antaranya pendekatan *two-pass* [7], dimana n-gram batas bawah digunakan dalam fase pembangkitan hipotesis, kemudian hipotesis K-terbaik dari hipotesis ini diberi skor ulang menggunakan model bahasa terdistribusi berskala besar. Hasil terjemahan yang dihasilkan terbukti dapat memperbaiki hipotesis yang dianggap terbaik oleh sistem tahap pertama. Emami dkk [8] melaporkan penelitiannya tentang *distributed language model* skala besar dalam konteks pengenalan suara dan mesin translasi. Mohaghegh [9] melaporkan peningkatan akurasi dengan meningkatkan peran model bahasa dalam mesin penerjemah English-Persian. Monz [10] melaporkan peningkatan akurasi dengan meningkatkan peran model bahasa dalam mesin penerjemah Arabic- dan Chinese-to-English. Maučec dkk. [11] melaporkan peningkatan akurasi dengan meningkatkan peran model bahasa dalam mesin penerjemah bahasa Serbian-Slovenian.

a. Model Bahasa

Model bahasa di desain untuk memperoleh probabilitas kemunculan kata-kata (atau token). Jika $W_1 = (w_1, \dots, w_L)$ menunjukkan *string token* L atas kosa kata terjemah. Model bahasa n-gram memberikan probabilitas untuk w_1^L menurut :

$$P(w_1^L) = \prod_{i=1}^L P(w_i | w_1^{i-1}) \approx \prod_{i=1}^L \hat{P}(w_i | w_{i-n+1}^{i-1}) \quad (3)$$

dimana aproksimasi tersebut mencerminkan asumsi Markov bahwa hanya n-1 token yang paling baru yang relevan saat memprediksi kata berikutnya.

Untuk setiap substring w, misalkan $f(w)$ menunjukkan frekuensi terjadinya substring tersebut pada deretan bahasa target yang ditentukan, biasanya sangat panjang yang disebut data pelatihan. Probabilitas maximum-likelihood (ML) untuk n-gram diberikan oleh frekuensi relatifnya.

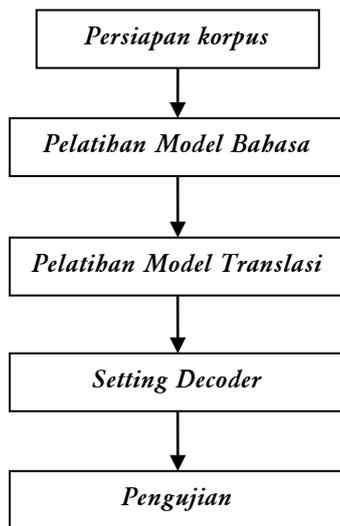
$$r(w_i | w_{i-n+1}^{i-1}) = \frac{f(w_{i-n+1}^i)}{f(w_{i-n+1}^{i-1})} \quad (4)$$

Pada prinsipnya, keakuratan prediktif model bahasa dapat ditingkatkan dengan meningkatkan orde n-gram. Namun, hal itu memperburuk masalah data yang jarang terjadi. Pekerjaan saat ini membahas tantangan dalam memproses suatu jumlah Data pelatihan cukup untuk model n-gram orde tinggi dan menyimpan dan mengelola nilai yang dihasilkan untuk penggunaan yang efisien oleh decoder.

2. Metode

2.1 Tahapan Eksperimen

Data pelatihan berupa korpus paralel bahasa Indonesia-Dayak Kanayatn yang bersumber dari hasil penelitian tentang tuning for quality untuk MPS bahasa Indonesia-Dayak Kanayatn [12]. Pada proses eksperimen, 667 kalimat digunakan untuk proses pengujian dan 3000 kalimat digunakan untuk proses training, sehingga total kalimat pada korpus yang digunakan adalah 3667 kalimat. Tahapan eksperimen yang dilakukan dapat dilihat pada Gambar 2. Persiapan korpus (preprocessing) dilakukan dengan melakukan proses cleaning, tokenisasi, dan lowercase terhadap korpus paralel yang sudah dipersiapkan. Model Bahasa yang digunakan pada sistem baseline adalah model trigram bahasa Dayak yang dilatih menggunakan toolkit SRILM [13], sedangkan korpus paralel yang siap digunakan kemudian dilatih untuk memperoleh word alignments, phrase table, language model, dan model combination weights. Base line yang digunakan pada riset ini dilatih menggunakan alat standar, yaitu GIZA ++ [14] untuk melatih penyelarasan kata dan Moses untuk pengodean berbasis frase. Moses adalah sebuah tools yang merupakan implementasi dari Mesin Penerjemah Statistik. Moses digunakan untuk melatih model statistik teks terjemahan dari bahasa sumber ke bahasa sasaran. Dalam menerjemahkan bahasa, Moses membutuhkan korpus dalam dua bahasa, bahasa sumber dan bahasa sasaran. Moses dirilis di bawah lisensi LGPL (Lesser General Public License) dan tersedia sebagai kode sumber dan binari untuk Windows dan Linux. Perkembangannya didukung oleh proyek EuroMatrix, dengan pendanaan oleh European Commission [15].



Gambar 2. Tahapan Eksperimen

rul , jadi dosen itu banyak pahitnya	rul , jadi dosen koa banyak pahitnya
pertama , mendidik mahasiswa tidak segampang yang kamu pikir	pamula' , nidik mahasiswa nana' sagampang nang kao mikir
kedua , sebagai dosen pegawai negeri , seberapa besar sih gajinya ? lupakan sementara bekerja di puskesmas daerah	kadua , jadi dosen pagawe nagari , sangahe aya' sih gajinya ? lupakan samintara pilihan bakaraja ka' puskesmas daerah
kalau boleh saya tahu , apa yang sedang kamu lakukan sekarang ini ? “ runut diuraikan pak arifin seperti anjuran bapak kepada anaknya	kade' mulih aku nau'an , ahe nang dah kao ngalakuan ampean ? “ runut dinguraian pak arifin lea anjuran apa' kada' kamuda nya
“ kebetulan saat ini ada yang menawari saya pinjaman modal dari bank sebesar rp 300 juta . “	“ kabatolan waktu nian ada nang nawari aku pinjaman modal dari bank samanyak rp 300 juta . “
kamu itu punya selera bisnis yang lain daripada yang lain , bakat juga ada , latihan sejak awal kuliah pun sudah kamu lakukan , apa lagi ? sudah , lupakan jadi dosen , biar yang lain yang mengurus masalah pendidikan	kao koa punya selera bisnis nang lain dari ka' nang lain , bakat uga' ada , latihan mulai' awal kuliah pun dah kao ngalakuan , ahe agi?? udah , lupa'atn jadi dosen , biar nang lain ngurusi masalah pendidikan
pertemuan kami sudahi	patamuan kami ngudahi
kian mantap langkah ini untuk memfokuskan diri sebagai pengusaha setelah diberikan semangat	kian mantap tongkang nian unto' mokuskan diri' jadi pangusaha laka' dimare' semangat sedemikian rupa
saat hubungan dengan anita sudah serius dan mendekati jenjang pernikahan , sempat saya perkenalkan kepada pak arifin di rumahnya	waktu hubungan man anita dah serius man nyamaki jenjang panganten , sampat aku nganalan ka' pak arifin ka' rumahnya
cukup lama anita diperhatikan pak arifin yang nampak sedikit kaget , “ lho , anita , kamu , kan , murid saya juga	lumayan lama anita dinelean pak arifin nang katelean sabebet takajut , “ lho , anita , kao , kan , murid ku uga'
kegagalan pertama saat usaha di luar kampus	kagagalan nang pamula'atn waktu usaha ka' luar kampus

Gambar 3. Potongan Korpus Indonesia-Dayak

Decoder sebagai mesin penerjemah selanjutnya di *setting* sesuai dengan strategi eksperimen yang dilakukan, yaitu dengan mengubah-ubah variabel *language model* yang digunakan. Untuk setiap *settingan*, dilakukan pengujian dengan input 667 kalimat yang telah dipersiapkan sebelumnya. Pengujian dilakukan dengan menggunakan metode evaluasi otomatis BLEU [16].

2.2 Strategi Eksperimen

Untuk melihat peran dari model bahasa yang digunakan pada MPS Indonesia-Dayak, pada eksperimen ini dilakukan empat strategi, yaitu :

- Language model* dilatih dari 3000 kalimat target korpus paralel, dengan kata lain kalimat referensi pengujian tidak diikutsertakan dalam pelatihan. Selanjutnya strategi pertama ini diuji untuk model 3-gram, 4-gram, 5-gram, 6-gram dan 7-gram.
- Language model* dilatih dari 3000 kalimat target korpus paralel ditambah 667 kalimat referensi, dengan kata lain kalimat referensi pengujian diikutsertakan dalam pelatihan. Selanjutnya strategi pertama ini diuji untuk model 3-gram, 4-gram, 5-gram, 6-gram dan 7-gram.
- Language model* dilatih dari 667 kalimat referensi dan 2333 kalimat target korpus paralel, kemudian ditambah 100 kalimat target korpus paralel yang tersisa untuk setiap eksperimen. Penambahan dilakukan sebanyak enam eksperimen, sehingga korpus yang digunakan masing-masing eksperimen adalah 3100, 3200, 3300, 3400, 3500, dan 3600 kalimat.
- Language model* dilatih dari 3000 kalimat target, kemudian ditambah 100 kalimat referensi untuk setiap eksperimen. Penambahan dilakukan sebanyak enam eksperimen, sehingga korpus yang digunakan masing-masing eksperimen adalah 3100, 3200, 3300, 3400, 3500, dan 3600 kalimat.

3. Hasil

Data pelatihan berupa korpus paralel bahasa Indonesia-Dayak (Gambar 3). Kolom sebelah kiri merupakan kumpulan kalimat dalam bahasa Indonesia, sedangkan kolom sebelah kanan merupakan kumpulan kalimat dalam bahasa Dayak, dimana setiap barisnya merupakan terjemahan dari kalimat yang bersesuaian. Potongan korpus tersebut merupakan potongan korpus yang sudah melewati tahapan proses *cleaning*, tokenisasi, dan *lowercase*.

Model bahasa dihasilkan dari proses pelatihan yang dilakukan terhadap bahasa target dari korpus paralel tersebut, yaitu bahasa Dayak. Sebagai *baseline* dilakukan pelatihan dengan model trigram (3-gram), kemudian dilakukan juga pelatihan untuk menghasilkan mesin pembandingan dengan model 4-gram, 5-gram, 6-gram dan 7-gram. Contoh model bahasa 3-gram dapat dilihat pada Gambar 3, sedangkan Contoh model bahasa 7-gram dapat dilihat pada Gambar 4. Dari hasil pelatihan, didapat jumlah pasangan token dengan probabilitasnya untuk masing-masing n-gram, yaitu : ngram 1=7830, ngram 2=32203, ngram 3=2186, ngram 4=597, ngram 5=185, ngram 6=67, dan ngram 7=26.

-0.7398544	man uber ka'
-0.9828925	, udah pasti
-0.7398544	ao' udah pak
-1.499522	nang udah ada
-1.016717	nang udah aku
-1.499522	nang udah ngantar
-1.499522	nang udah tuha
-1.202949	aku uga' minta
-1.499522	aku uga' pancah
-0.4388244	diri' uga' harus
-0.6963823	koa uga' aku
-0.4388244	ku uga' </s>
-0.6149157	nana' uga' </s>
-1.313886	tapi uga' ka'
-0.9159456	ka' ui ,

Gambar 4. Potongan Model Bahasa 3-gram

-0.1760913	ka' babah chairul tanjung foundation (ctf
-0.1760913	jadi ketua komite ekonomi nasional (ken
-0.1760913	rumah anak madani (ram) nang
-0.1760913	<s> dua tahun laka' koa , ka'
-0.1760913	london school of public relation , jakarta
-0.1760913	, baik umat islam maopun agama lain
-0.1760913	ka' gang abu , maraga batutulis ,
-0.1760913	babah chairul tanjung foundation (ctf)
-0.1760913	nang labih edo' ka' masa depan </s>
-0.1760913	man pandapatan per kapita 18.000 dollar as
-0.1760913	school of public relation , jakarta </s>
-0.1760913	idup ka' dunia nian , jukut koa
-0.1760913	jadi bank ta'edo' man aset ka'babah rp
-0.1760913	bank mega syariah surabaya man kapala
-0.1760913	cabang
-0.1760913	ketua komite ekonomi nasional (ken)
-0.1760913	baik umat islam maopun agama lain </s>
-0.1760913	cabang bank mega syariah surabaya man
-0.1760913	kapala

Gambar 5. Potongan Model Bahasa 7-gram

Tabel 1. Hasil pengujian strategi 1

Mesin	n-gram	Nilai BLEU (%)
1A	3-gram	45.56
1B	4-gram	45.57
1C	5-gram	45.60
1D	6-gram	45.60
1E	7-gram	45.60

Tabel 2. Hasil pengujian strategi 2

Mesin	n-gram	Nilai BLEU (%)
2A	3-gram	56.21
2B	4-gram	56.24
2C	5-gram	56.21
2D	6-gram	56.15
2E	7-gram	56.18

Eksperimen strategi pertama menggunakan korpus paralel sebesar 3000 kalimat dan korpus monolingual juga sebesar 3000 kalimat, hasil eksperimen yang dihasilkan dapat dilihat pada Tabel 1. Hasil eksperimen menunjukkan bahwa dengan penambahan n-gram pada korpus monolingual yang diambil dari korpus paralel tidak menunjukkan peningkatan akurasi (diwakili dengan nilai BLEU) yang signifikan. Nilai tertinggi pada model 7-gram hanya dapat meningkatkan akurasi sebesar $((45,60-45,56)/45,56) * 100\% = 0,09\%$.

Tabel 3. Hasil pengujian strategi 3

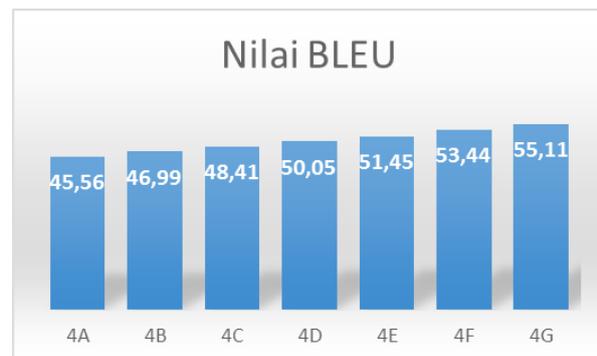
Mesin	Korpus monolingual	Nilai BLEU (%)
3A	3000	56.59
3B	3100	56.63
3C	3200	56.62
3D	3300	56.61
3E	3400	56.58
3F	3500	56.52
3G	3600	56.45

Tabel 4. Hasil pengujian strategi 4

Mesin	Korpus monolingual	Nilai BLEU (%)
4A	3000	45.56
4B	3100	46.99
4C	3200	48.41
4D	3300	50.05
4E	3400	51.45
4F	3500	53.44
4G	3600	55.11

Tabel 5. Peningkatan Akurasi pada strategi 4

Mesin	Nilai BLEU (%)	Peningkatan Akurasi (%)
4A	45.56	
4B	46.99	3.14
4C	48.41	3.02
4D	50.05	3.39
4E	51.45	2.80
4F	53.44	3.87
4G	55.11	3.13



Gambar 6. Grafik Hasil Pengujian Strategi 4

Eksperimen strategi kedua menggunakan korpus paralel sebesar 3000 kalimat dan korpus monolingual sebesar 3667 kalimat, hasil eksperimen yang dihasilkan dapat dilihat pada Tabel 2. Hasil eksperimen menunjukkan bahwa dengan penambahan n-gram pada korpus monolingual yang diambil dari korpus paralel tidak menunjukkan peningkatan akurasi (diwakili dengan nilai BLEU) yang signifikan, bahkan pada model 6-gram lebih rendah dari baseline 3-gram. Nilai tertinggi pada model 4-gram hanya dapat meningkatkan akurasi sebesar $((56,24-56,21)/56,21) * 100\% = 0,05\%$.

Eksperimen strategi ketiga menggunakan korpus paralel sebesar 3000 kalimat dan korpus monolingual sebesar 3000 sampai 3600 kalimat, 3000 kalimat baseline terdiri dari 667 kalimat referensi dan 2333 kalimat target korpus paralel. Hasil eksperimen yang dihasilkan dapat dilihat pada Tabel 3 menunjukkan bahwa dengan penambahan kuantitas korpus monolingual yang diambil dari korpus paralel tidak menunjukkan peningkatan akurasi yang signifikan, bahkan menurun dari mesin 3B sampai 3G. Nilai tertinggi pada mesin 3B hanya dapat meningkatkan akurasi sebesar $((56,63-56,59)/56,59) * 100\% = 0,07\%$.

Eksperimen strategi keempat menggunakan korpus paralel sebesar 3000 kalimat dan korpus monolingual sebesar 3000 sampai 3600 kalimat, 3000 kalimat baseline keseluruhannya diambil dari kalimat target korpus paralel. Hasil eksperimen yang dihasilkan dapat dilihat pada Tabel 4 menunjukkan bahwa dengan penambahan kuantitas korpus monolingual yang diambil dari kalimat referensi menunjukkan peningkatan akurasi yang signifikan. Nilai tertinggi pada mesin 3G dengan 3600 korpus monolingual dapat meningkatkan akurasi sebesar $((55,11-45,56)/45,56) * 100\% = 20,96\%$.

4. Diskusi

Eksperimen yang dilakukan pada strategi 1 dan 2 menunjukkan bahwa penggunaan model n-gram dari 3-gram sampai 7-gram tidak berpengaruh terhadap akurasi mesin penerjemah Indonesia-Dayak dengan korpus paralel sebesar 3000 kalimat. Hal ini disebabkan terlalu kecilnya kuantitas kalimat yang digunakan pada korpus. Kecilnya kuantitas kalimat korpus mengakibatkan tidak bervariasinya probabilitas setiap pasangan token seperti terlihat pada model bahasa 7-gram pada gambar 4, dengan demikian, untuk MPS yang menggunakan korpus kecil sebaiknya tetap menggunakan model bahasa 3-gram.

Eksperimen penambahan kuantitas korpus monolingual yang digunakan untuk pelatihan model bahasa secara bertahap, hasilnya diperlihatkan oleh strategi 3 dan 4. Dari hasil eksperimen didapatkan hasil bahwa hasil terbaik didapatkan dengan menambah kuantitas korpus monolingual di luar korpus paralel pada strategi 4 (lihat Gambar 6), dengan kata lain, korpus monolingual diambil dari bahasa target korpus paralel, selanjutnya ditambahkan dengan kalimat lain di luar kalimat yang ada pada korpus paralel.

Hasil eksperimen pada strategi 4 menunjukkan peningkatan yang cukup signifikan untuk setiap penambahan 100 kalimat korpus monolingual seperti terlihat pada Gambar 6 dan Tabel 5.

Dari hasil penelitian ini, didapat kesimpulan bahwa peran model bahasa cukup penting dalam mengantisipasi kalimat-kalimat yang akan diterjemahkan pada MPS, terutama saat frase di dalam kalimat tersebut tidak terdapat di dalam model translasi. Hal ini tentu akan lebih dirasakan pada MPS dengan sumberdaya kecil, karena kemungkinan kalimat yang akan diterjemahkan tidak terdapat pada model translasi tentu sangat besar dibandingkan dengan MPS dengan sumber daya besar.

5. Kesimpulan

Penggunaan model n-gram dari 3-gram sampai 7-gram tidak berpengaruh terhadap akurasi mesin penerjemah Indonesia-Dayak, disarankan untuk MPS yang menggunakan korpus kecil sebaiknya tetap menggunakan model bahasa 3-gram.

Hasil terbaik untuk meningkatkan peran model bahasa adalah dengan menggunakan bahasa target korpus paralel sebagai korpus monolingual, selanjutnya ditambahkan sebanyak-banyaknya dengan kalimat lain di luar kalimat yang ada pada korpus paralel. Kedepannya, perlu dilakukan penelitian dengan metode yang sama untuk bahasa-bahasa daerah lainnya yang ada di Indonesia.

6. Daftar Pustaka

- [1] T. McEnery, *Corpus-Based Language Studies: An Advanced Resource*, 2006.
- [2] K. Kirchhoff dan M. Yang, "Improved Language Modeling for Statistical Machine Translation," dalam *Proceedings of the ACL Workshop on Building and Using Parallel Texts*, Michigan, 2005.
- [3] H. Sujaini dan B. P. Arif, "Strategi Memperbaiki Kualitas Korpus untuk Meningkatkan Kualitas Mesin Penerjemah Statistik," dalam *Seminar Nasional Teknologi Informasi XI*, Jakarta, 2014.
- [4] S. Mandira, H. Sujaini dan B. P. Arif, "Perbaikan Probabilitas Lexical Model untuk Meningkatkan Akurasi Mesin Penerjemah Statistik," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, Vol. 2, No. 1, 2016.
- [5] Y. Jarob, H. Sujaini dan N. Safriadi, "Uji Akurasi Penerjemahan Bahasa Indonesia – Dayak Taman dengan Penandaan Kata Dasar dan Imbuhan," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, Vol. 2, No. 2, 2016.
- [6] R.A. Nugroho, T.B. Adji, dan B.S. Hantono, *Penerjemahan Bahasa Indonesia dan Bahasa Jawa Menggunakan Metode Statistik Berbasis Frasa*, dalam *Seminar Nasional Teknologi Informasi dan Komunikasi 2015 (SENTIKA 2015)*, Yogyakarta, 2015.
- [7] Y. Zhang, A. S. Hildebrand dan S. Vogt, "Distributed language modeling for n-best list," dalam *Proceedings of EMNLP-2006*, Sydney, 2006.
- [8] A. Emami, K. Papineni dan J. Sorensen, "Large-scale distributed language modeling," dalam *Proceedings of ICASSP-2007*, Honolulu, 2007.
- [9] M. Mohaghegh, A. Sarrafzadeh dan T. Moir, "Improved Language Modeling for English-Persian Statistical Machine," dalam *SSST-4, Fourth Workshop on Syntax and Structure in Statistical Translation*, Beijing, 2010.

- [10] C. Monz, "Statistical Machine Translation with Local Language Models," dalam *Conference on Empirical Methods in Natural Language Processing*, Edinburgh, 2011.
- [11] M. S. Maučec, G. Donaj dan Z. Kacič, "Improving statistical machine translation with additional language models," dalam *Human Language Technologies As a Challenge for Computer Science and Linguistics*, Poznań, 2013.
- [12] M. Hasbiansyah, H. Sujaini dan N. Safriadi, "Tuning For Quality untuk Uji Akurasi Mesin Penerjemah Statistik (MPS) Bahasa Indonesia - Bahasa Dayak Kanayatn", *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 4, No. 1, 2016.
- [13] A. Stolcke, J. Zheng, W. Wang dan V. Abrash, "SRILM at sixteen: Update and outlook," dalam *Automatic Speech Recognition and Understanding (ASRU), 2011 IEEE Workshop*, Waikoloa, 2011.
- [14] F. J. Och dan H. Ney, "A Systematic Comparison of Various Statistical Alignment Models," *Computational Linguistics*, vol. 1, no. 29, pp. 19-51, 2003.
- [15] P. Koehn, *Statistical machine translation*, New York: Cambridge University Press, 2010.
- [16] K. Papineni, S. Roukos, T. Ward dan W.-J. Zhu, "BLEU: A Method For Automatic Evaluation of Machine Translation," dalam *Proceedings of the 40th Annual Meeting of the Association of Computational Linguistics (ACL)*, Pennsylvania, 2002.