

Analisis Penggabungan Korpus dari Hadits Nabi dan Alquran untuk Mesin Penerjemah Statistik

Hafidz Ardhi^{#1}, Herry Sujaini^{*2}, Arif Bijaksana Putra Negara^{#3}

[#] Program Studi Teknik Informatika Universitas Tanjungpura
Pontianak

¹hafidzardhi08@gmail.com

²herry.sujaini@ee.untan.ac.id

³arifbpn@gmail.com

Abstract — Each region has different language to communicate. A communication can run well if each other can understand the language that use in communication process. Machine translation is an automatic translation machine to translate a text from a language to another language. In the machine translation there will be an automatic evaluation. Automatic evaluation is needed to measure the quality of translation text from machine translation using automatic metric. The metric is use to determine score toward quality in various ways until get percentage at the final result. Evaluation of translation machine system using automatic metric is quick, easy and inexpensive way rather than human evaluation. BLEU is a common metric used by researcher to evaluate machine translation. For this research, researcher used Arabic Language. Corpus that used are corpus of Al-quran, corpus of hadith, and combined corpus. The corpus will be tested with the type of sentence and 4 level numbers of sentences. The test will be done in two times. First, test without MADAMIRA. Second, test using MADAMIRA. The result of tested without MADAMIRA produce BLEU score for corpus of Al-quran in amount of 10,56%, corpus of hadith 27,65%, and combined corpus 15,41%. In the other hand, the result of tested corpus used MADAMIRA got result of BLEU for corpus of Al-quran 1,44%, corpus of hadith 32,90% and combined corpus 41,46%.

Keyword : Statistical machine translation, Automatic Evaluation, , Hadits, Alqur'an, MADAMIRA.

Abstrak— Setiap daerah memiliki bahasa berbeda-beda yang digunakan dalam berkomunikasi. Suatu komunikasi dapat berjalan dengan baik jika kedua pihak yang terlibat dalam komunikasi, mengerti dengan bahasa yang digunakan selama proses interaksi berlangsung. Mesin penerjemah merupakan alat penerjemah otomatis pada sebuah teks dari satu bahasa ke bahasa lainnya. Pada mesin penerjemah akan dilakukan proses evaluasi yang dilakukan secara otomatis. Evaluasi secara otomatis diperlukan untuk mengukur kualitas dari terjemahan mesin penerjemah dengan

menggunakan automatic metric. BLEU adalah metric yang paling sering digunakan peneliti untuk mengevaluasi mesin penerjemah. Pada penelitian ini peneliti menggunakan bahasa Arab. Korpus yang digunakan adalah korpus alquran, korpus hadits, dan korpus gabungan. Korpus akan diuji dengan tipe kalimat dan 4 tingkat jumlah kalimat. Pengujian akan dilakukan dua kali. Pertama, pengujian terhadap korpus tanpa menggunakan MADAMIRA. Kedua, pengujian terhadap korpus menggunakan MADAMIRA. Hasil dari pengujian terhadap korpus tanpa menggunakan MADAMIRA menghasilkan nilai BLEU pada korpus alquran sebesar 10,56 %, korpus hadits sebesar 27,65%, dan korpus gabungan sebesar 15,41%. Sedangkan hasil dari pengujian yang dilakukan terhadap korpus yang menggunakan MADAMIRA diperoleh nilai BLEU pada korpus alquran sebesar 1,44%, korpus hadits sebesar 32,90%, dan korpus gabungan sebesar 41,46%.

Kata kunci : Mesin penerjemah statistik, Evaluasi otomatis, Hadits, Alquran, MADAMIRA

I. PENDAHULUAN

Mesin penerjemah (MP) merupakan salah satu teknologi yang saat ini banyak dikembangkan. Salah satunya adalah mesin penerjemah statistik. Mesin penerjemah merupakan alat penerjemah otomatis pada sebuah teks dari satu bahasa ke bahasa lainnya. Pengembangan mesin penerjemah membutuhkan korpus. Korpus merupakan data-data yang dipakai sebagai sumber bahan untuk menerjemahkan suatu bahasa. Korpus yang digunakan pada penelitian ini adalah korpus paralel. Korpus paralel berisikan seperangkat teks terjemahan dalam satu bahasa dan teks sumbernya dalam bahasa yang berbeda.

Beberapa penelitian sebelumnya untuk MP bahasa Arab telah dilakukan, diantaranya adalah untuk bahasa

Inggris-Arab yang secara lengkap telah disusun state of the art-nya[1], sedangkan untuk bahasa Indonesia, beberapa penelitian terakhir yang sudah dilakukan diantaranya adalah penelitian perbaikan probabilitas lexical model untuk meningkatkan akurasi MPS bahasa Indonesia-Jawa [2], penelitian uji akurasi MPS bahasa Indonesia-Dayak Taman dengan penandaan kata dasar dan imbuhan [3], dan penelitian terhadap MPS bahasa Indonesia-Jawa yang menunjukkan hasil evaluasi 44,02 % untuk Indonesia-Jawa dan 48,77 % untuk Jawa-Indonesia [4].

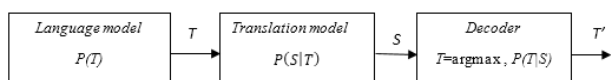
Pada penelitian ini korpus paralel yang digunakan adalah bahasa Arab dan bahasa Indonesia. Sumber korpus yang digunakan bersumber dari Alquran dan Hadits. Paper ini berisi analisis terhadap nilai akurasi mesin penerjemah statistik terhadap korpus paralel bahasa Arab dan bahasa Indonesia yang bersumber dari alquran dan hadits nabi. Mesin penerjemah statistik akan menggunakan metode BLEU, untuk menguji tingkat akurasi terjemahan.

A. Mesin Penerjemah Statistik (NLP)

Statistical machine translation atau mesin penerjemah statistik merupakan salah satu jenis mesin penerjemah dengan menggunakan pendekatan statistik. Pendekatan statistik yang digunakan adalah konsep probabilitas. Setiap pasangan kalimat (S,T) akan diberikan sebuah P(T|S) yang diinterpretasikan sebagai distribusi probabilitas dimana sebuah penerjemah akan menghasilkan T dalam bahasa tujuan ketika diberikan S dalam bahasa sumber. Sebagai contoh, proses penerjemahan dari kalimat Bahasa Prancis ke Bahasa Inggris dapat diinterpretasikan dengan P(e|f) . P(e|f) merupakan distribusi probabilitas kalimat e (Bahasa Inggris) terhadap kalimat f (Bahasa Perancis). Dengan menggunakan teorema Bayes pada P(e|f) , didapatkan:

$$P(T/S) = \frac{P(S|T) \cdot P(T)}{P(S)} \tag{1}$$

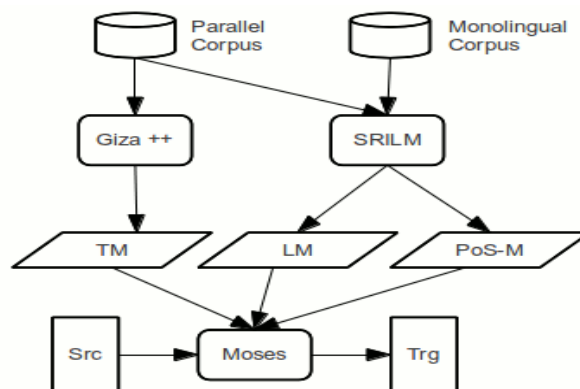
Menurut Christopher D Manning dan Hinrich Schutze, dalam mesin penerjemahan berbasis statistik terdapat tiga buah komponen yang terlibat dalam proses penerjemahan kalimat dari suatu bahasa ke bahasa lain, yaitu: *language model*, *translation model*, dan *decoder* [5].



Gambar 1 Komponen Statistical Machine Translation

Penjelasan singkat dari gambar 1 adalah, *language model* menghasilkan kalimat bahasa T, *translation model* mengirimkan kalimat bahasa T sebagai kalimat bahasa S.

Decoder mencari kalimat bahasa T' yang paling mungkin yang telah menimbulkan kalimat S. Secara umum, arsitektur mesin penerjemah statistik Moses ditunjukkan pada Gambar 2 .



Gambar 2 Arsitektur Mesin Penerjemah Statistik Moses [6]

Sumber data utama yang dipergunakan adalah *parallel corpus* dan *monolingual corpus*. Proses *training* terhadap *parallel corpus* menggunakan GIZA++ menghasilkan *translation model* (TM). Proses *training* terhadap bahasa target pada *parallel corpus* ditambah dengan *monolingual corpus* bahasa target menggunakan SRILM menghasilkan *language model* (LM), sedangkan *PoS model* (PoS-M) dihasilkan dari bahasa target pada *parallel corpus* yang setiap katanya sudah ditandai dengan PoS. *TM*, *LM* dan *PoS-M* digunakan untuk menghasilkan *decoder Moses*. Selanjutnya *Moses* digunakan sebagai mesin penerjemah untuk menghasilkan bahasa target dari *input* kalimat dalam bahasa sumber.

Moses adalah salah satu Mesin Penerjemah Statistik yang memungkinkan untuk menerjemahkan secara otomatis setiap pasangan bahasa. *Moses* digunakan untuk melatih model statistic teks terjemahan dari bahasa sumber ke bahasa sasaran. Saat melakukan penerjemahan Bahasa [2]. *Moses* membutuhkan korpus dua bahasa, bahasa sumber dan bahasa sasaran. *Moses* dirilis di bawah lisensi LGPL (Lesser General Public License) dan tersedia sebagai kode sumber dan binary untuk Windows dan Linux. Perkembangannya didukung oleh proyek EuroMatrix, dengan pendanaan oleh European Commission.

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. Korpus dapat diklasifikasikan ke dalam delapan jenis, yaitu korpus khusus (*specialized corpus*), korpus umum (*general corpus*), korpus komparatif (*comparable corpus*), korpus paralel (*parallel corpus*), korpus pemelajar (*learner corpus*), korpus pedagogis (*pedagogic corpus*), korpus historis atau korpus diakronis (historical or diachronic corpus), dan korpus monitor (*monitor corpus*). [7].

BLEU (*Bilingual Evaluation Understudy*) adalah sebuah algoritma yang berfungsi untuk mengevaluasi kualitas dari sebuah hasil terjemahan yang telah diterjemahkan oleh mesin dari satu bahasa alami ke bahasa lain. Ide utama dibalik ini adalah “semakin dekat terjemahan sebuah mesin dengan terjemahan manusia maka akan semakin baik”. BLEU mengukur *modified n-gram precision score* antara hasil terjemahan otomatis dengan terjemahan rujukan dan menggunakan konstanta yang dinamakan *brevity penalty*. [8].

Cara sederhana untuk menghitung *precision* adalah dengan menghitung jumlah kata pada hasil terjemahan (*unigrams*) yang cocok dengan rujukan dan dibagi dengan total jumlah kata (*unigrams*) yang ada pada hasil terjemahan. Namun cara tersebut memiliki beberapa kelemahan. *Stastical machine translation* dapat menghasilkan dengan berlebihan kata-kata yang benar yang mengakibatkan dihasilkannya terjemahan yang mustahil namun memiliki *precision* tinggi. Dengan demikian sebuah kata rujukan harus dianggap telah muncul setelah sebuah kata hasil terjemahan yang cocok teridentifikasi. Untuk mencapai tujuan tersebut maka Papineni melakukan beberapa perubahan dan dihasilkanlah metode yang dinamakan *modified n-gram precision*. [9].

Brevity penalty digunakan untuk mencegah kalimat pendek memperoleh nilai yang tinggi. Dengan kata lain, BLEU diperoleh dari hasil perkalian antara *brevity penalty* dengan rata-rata geometri dari *modified precision score*. Jadi agar dapat menghasilkan nilai yang tinggi dalam BLEU, panjang kalimat hasil terjemahan harus mendekati panjang kalimat referensi dan kalimat hasil terjemahan harus dapat menghasilkan kata serta urutan yang sama dengan kalimat referensi. Nilai dari BLEU berada dalam rentang 0 sampai 1. Semakin tinggi nilai BLEU, maka akan semakin akurat dengan rujukan. Rumus menghitung BLEU adalah sebagai berikut [8]:

$$BP_{BLEU} = \begin{cases} 1 & \text{if } c > r \\ e^{(1-r/c)} & \text{if } c \leq r \end{cases} \quad (1)$$

$$P_n = \frac{\sum_{C \in \text{corpus}} \sum_{n\text{-gram} \in C} \sum \text{count}_{clip(n\text{-gram})}}{\sum_{C \in \text{corpus}} \sum_{n\text{-gram} \in C} \sum \text{count}(n\text{-gram})} \quad (2)$$

$$BLEU = BP_{BLEU} \cdot e^{\sum_{n=1}^N w_n \log p_n} \quad (3)$$

Keterangan :

BP = *brevity penalty*

c = jumlah kata dari hasil terjemahan otomatis

r = jumlah kata rujukan

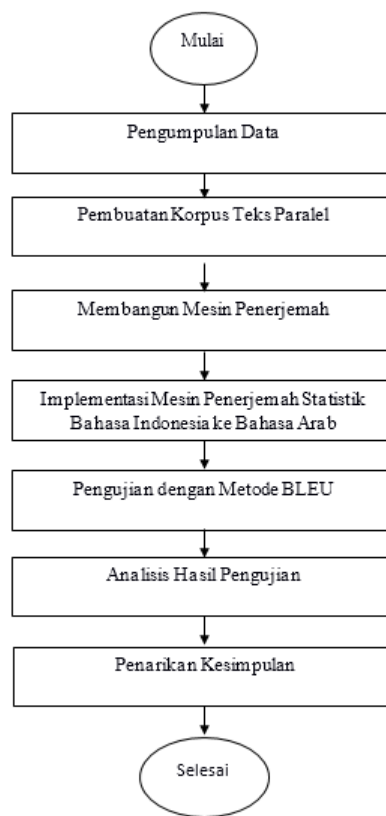
P_n = *modified precision score*

w_n = 1/N (standar nilai N untuk BLEU adalah 4)

p_n = jumlah n-gram hasil terjemahan yang sesuai dengan rujukan dibagi jumlah n-gram hasil terjemahan.

II. METODOLOGI

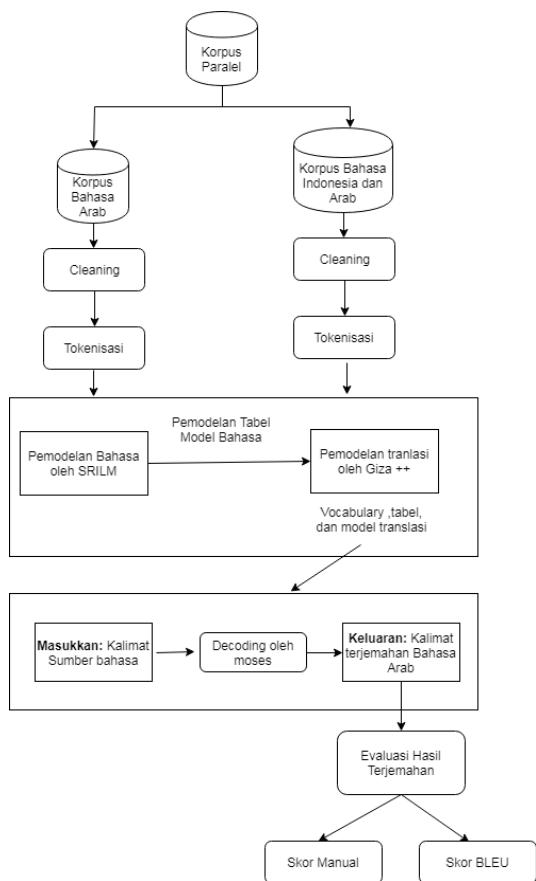
Metodologi penelitian yang dilakukan dijelaskan melalui alir diagram penelitian pada gambar 3 : Data penelitian yang digunakan berupa kitab Bulughul Maram dan Bukhari Muslim. Kitab tersebut selanjutnya diolah menjadi korpus teks paralel bahasa Arab dan bahasa Indonesia. Jumlah korpus alquran adalah 6.236, kalimat, korpus hadits berjumlah 6.887 kalimat, sedangkan korpus gabungan berjumlah 13.122 kalimat.



Gambar 3 Diagram Penelitian

III. HASIL DAN DISKUSI

Pada penelitian ini, perancangan arsitektur sistem terdiri dari beberapa proses yaitu tahap awal proses persiapan data, pemodelan bahasa, decoding dan evaluasi hasil terjemahan. Proses - proses tersebut dapat dilihat pada gambar 4



Gambar 4 Arsitektur Mesin Penerjemah Statistik dari Bahasa Indonesia ke Bahasa Arab

B. Pengujian

Pengujian akan dilakukan dua kali. Pertama, pengujian terhadap korpus yang proses tokenisasi tanpa menggunakan MADAMIRA. Kedua pengujian terhadap korpus yang proses tokenisasi menggunakan MADAMIRA.

1. Pengujian Menggunakan 3 tipe kalimat..

Pengujian menggunakan 3 jenis kalimat, yaitu kalimat tanya, kalimat perintah, dan kalimat Narasi. Ketiga jenis kalimat masing – masing akan diujikan pada korpus hadits nabi, alquran dan korpus gabungan antara hadits nabi dan alquran. Pengujian dilakukan dua kali terhadap ketiga korpus tersebut. Hasil dari pengujian dengan 3 tipe kalimat tersebut dapat dilihat pada tabel I dan tabel II.

TABEL I

HASIL PENGUJIAN MENGGUNAKAN 3 TIPE KALIMAT TERHADAP KORPUS TANPA MENGGUNAKAN MADAMIRA

Jenis Kalimat	Nilai BLEU Alquran	Nilai BLEU Hadits	Nilai BLEU Gabungan
Kalimat Narasi	10%	0 %	0%
Kalimat Perintah	10%	0%	0%
Kalimat Tanya	0%	0%	0%

TABEL III

HASIL PENGUJIAN MENGGUNAKAN 3 TIPE KALIMAT TERHADAP KORPUS MENGGUNAKAN MADAMIRA

Jenis Kalimat	Nilai BLEU Alquran	Nilai BLEU Hadits	Nilai BLEU Gabungan
Kalimat Narasi	0%	0 %	0%
Kalimat Perintah	0%	0%	0%
Kalimat Tanya	0%	0%	0,5%

2. Pengujian 4 Tingkat Jumlah Kalimat

Pengujian dilakukan dengan tingkatan jumlah kalimat sebesar 2000 kalimat,3000 kalimat, 4000 kalimat, dan 5000 kalimat. Korpus hadits, korpus alquran dan korpus gabungan akan diujikan dengan keempat tingkat jumlah kalimat.

a. Pengujian 4 Tingkat Jumlah Kalimat Terhadap Korpus Tanpa Menggunakan MADAMIRA

TABEL III

PENGUJIAN 4 TINGKAT JUMLAH KALIMAT TERHADAP KORPUS ALQURAN

No	Jumlah Kalimat	Nilai BLEU
1.	2000 kalimat	7,67 %
2.	3000 kalimat	7,45 %
3.	4000 kalimat	6,77 %
4.	5000 kalimat	5,97 %

Tabel III menunjukkan korpus alquran mengalami penurunan skor BLEU pada setiap 4 tingkat jumlah kalimat. Penurunan skor BLEU terhadap korpus AlQuran pada tingkat pertama sampai tingkat keempat jumlah kalimat sebesar 1,7 %.

TABEL IV

PENGUJIAN 4 TINGKAT JUMLAH KALIMAT TERHADAP KORPUS HADITS

No	Jumlah Kalimat	Nilai BLEU
1.	2000 kalimat	21,37%
2.	3000 kalimat	21,24%
3.	4000 kalimat	19,13%
4.	5000 kalimat	17,72%

Tabel IV menunjukkan korpus hadits juga mengalami penurunan nilai BLEU pada pengujian 4 tingkat, seperti halnya pada korpus alquran.

TABEL V

PENGUJIAN 4 TINGKAT JUMLAH KALIMAT TERHADAP KORPUS GABUNGAN

No	Jumlah Kalimat	Nilai BLEU
1.	2000 kalimat	13,02 %
2.	3000 kalimat	12,13 %
3.	4000 kalimat	11,35 %
4.	5000 kalimat	10,32 %

Tabel V menunjukan penurun nilai bleu pada hasil pengujian korpus gabungan yang cukup besar dibandingkan dengan korpus alquran. Pada korpus gabungan penurunan yang terjadi sebesar 2,7 %..

b. Pengujian 4 Tingkat Jumlah Kalimat Terhadap Korpus yang Menggunakan MADAMIRA

TABEL VI

PENGUJIAN 4 TINGKAT JUMLAH KALIMAT TERHADAP KORPUS ALQURAN YANG MENGGUNAKAN MADAMIRA

No	Jumlah Kalimat	Nilai BLEU
1.	2000 kalimat	12,56 %
2.	3000 kalimat	12,38 %
3.	4000 kalimat	12,59 %
No	Jumlah Kalimat	Nilai BLEU
4.	5000 kalimat	12,32 %

Tabel VI memperlihatkan pada pengujian ini terjadi kenaikan dan penurunan. Pada 2000 kalimat, nilai BLEU yang dihasilkan sebesar 12,56% sedangkan pengujian dengan menggunakan 3000 kalimat, nilai BLEU yang dihasilkan 12,38%. Hal ini memperlihatkan terjadinya penurunan nilai BLEU akan tetapi pada pengujian menggunakan 4000 kalimat terjadi kenaikan. Nilai BLEU yang dihasilkan pada 4000 kalimat sebesar 12,59 % dan terjadi penurunan lagi pada pengujian menggunakan 5000 kalimat. Penurunan nilai BLEU sebesar 0,27 %.

TABEL VII

PENGUJIAN 4 TINGKAT JUMLAH KALIMAT TERHADAP KORPUS HADITS YANG MENGGUNAKAN MADAMIRA

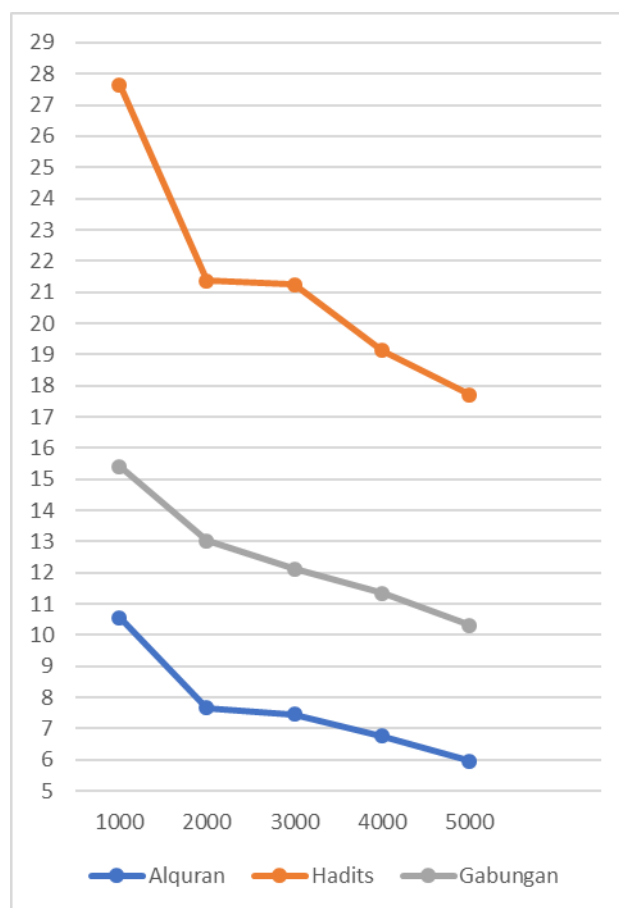
No	Jumlah Kalimat	Nilai BLEU
1.	2000 kalimat	12,56 %
2.	3000 kalimat	12,38 %
3.	4000 kalimat	12,59 %
4.	5000 kalimat	12,32 %

Pada Tabel VII memperlihatkan nilai BLEU tertinggi dihasilkan dengan menggunakan kalimat berjumlah 3000. Pengujian dengan menggunakan 5000 kalimat memiliki nilai BLEU paling kecil diantara ketiga jumlah kalimat yang lain.

TABEL VIII

PENGUJIAN 4 TINGKAT JUMLAH KALIMAT TERHADAP KORPUS GABUNGAN YANG MENGGUNAKAN MADAMIRA

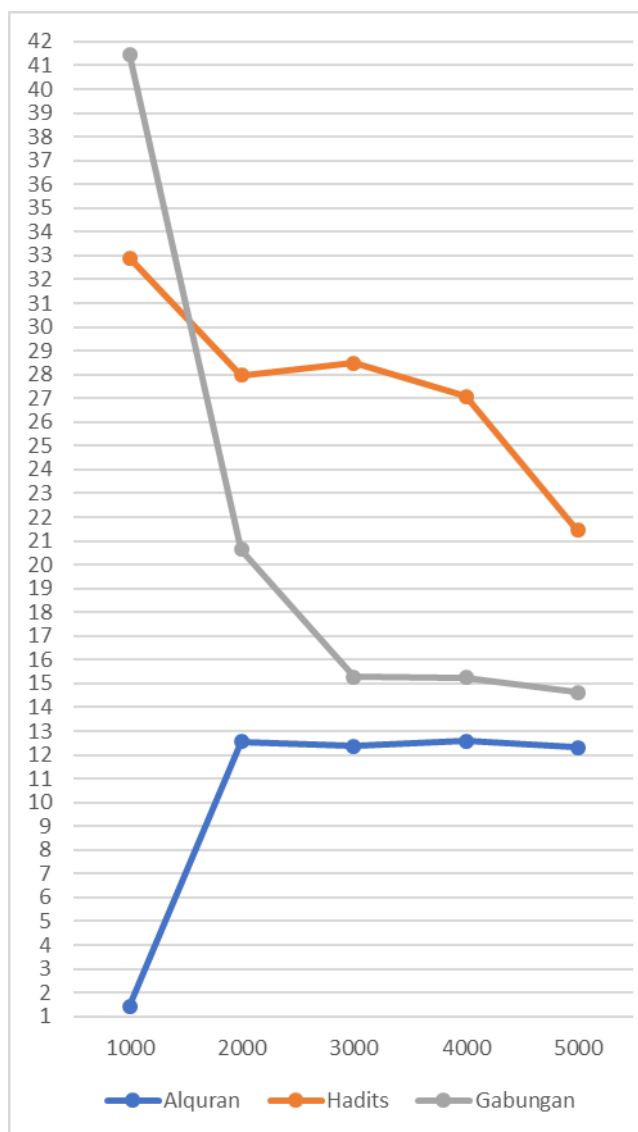
No	Jumlah Kalimat	Nilai BLEU
1.	2000 kalimat	20,65 %
2.	3000 kalimat	15,28 %
3.	4000 kalimat	15,26 %
4.	5000 kalimat	14,62 %



Gambar 5 Grafik Pengujian Terhadap Korpus Tanpa Menggunakan MADAMIRA

Tabel VIII menunjukkan semakin banyak jumlah kalimat pada korpus gabungan menghasilkan nilai BLEU yang rendah. Pada 2000 kalimat menghasilkan nilai BLEU sebesar 20,65 % sedangkan pada 3000 kalimat menghasilkan nilai BLEU sebesar 15,28 %. Hal ini menunjukkan terjadinya penurunan nilai BLEU. Adapun grafik dari percobaan pengujian terhadap korpus tanpa menggunakan MADAMIRA maupun korpus menggunakan MADAMIRA.

Gambar 5 menunjukkan korpus yang tanpa menggunakan MADAMIRA cenderung mengalami penurunan nilai BLEU.



Gambar 6 Grafik Pengujian Terhadap Korpus Menggunakan MADAMIRA

Pada gambar 6 dapat dilihat penurunan nilai BLEU terjadi pada korpus gabungan dan hadits. Korpus alquran memiliki nilai BLEU yang tidak jauh beda antara 1000 kalimat sampai dengan 5000 kalimat.

3. Pengujian oleh Ahli Bahasa

Pengujian oleh ahli bahasa juga dilakukan dua kali. Adapun hasil penilaian ahli bahasa digunakan 100 kalimat sebagai sample. Nilai dari ahli bahasa terhadap korpus tanpa menggunakan MADAMIRA dan korpus yang menggunakan MADAMIRA dapat dilihat pada tabel IX dan tabel X.

TABEL IX

PENILAIAN AHLI BAHASA TERHADAP KORPUS TANPA MENGGUNAKAN MADAMIRA

No	Nilai AKurasi	Jumlah
1.	100%	11 kalimat
2.	90 %	6 kalimat
3.	80%	2 kalimat
4.	0 %	81 kalimat

Tabel IX menunjukkan dari 100 kalimat yang dinilai oleh bahasa. Kalimat yang mempunyai presentase akurasi 100 % sebanyak 11 kalimat sedangkan yang mempunyai nilai akurasi 0 % sebanyak 81 kalimat.

TABEL X

PENILAIAN AHLI BAHASA TERHADAP KORPUS YANG MENGGUNAKAN MADAMIRA

No	Nilai AKurasi	Jumlah
1.	100%	0 kalimat
2.	90 %	10 kalimat
3.	80%	10 kalimat
4.	0 %	2 kalimat

Tabel X menunjukkan bahwa tidak ada kalimat yang mendapat nilai 100 %. Hal ini menunjukkan tidak ada satupun kalimat terjemahan yang dihasilkan tepat dengan bahasa target. Adapun kriteria yang menjadi dasar penilaian oleh ahli bahasa terhadap kalimat – kalimat hasil terjemahan adalah :

- Penyambungan huruf-huruf arab yang terdapat pada kalimat.
- Penulisan kata ganti.
- Tatabahasa yang terdapat pada kalimat

C. Analisis Hasil Pengujian

Dari pengujian dengan menghitung akurasi kecocokan frasa jeda satu kalimat penuh dan perhitungan precision & recall serta *F-measure* yang telah dilakukan, analisis yang dapat disimpulkan sebagai berikut:

- Nilai BLEU terhadap korpus yang tidak menggunakan MADAMIRA pada korpus alquran sebesar 10,56%, korpus hadits sebesar 27,65 %, dan korpus gabungan sebesar 15,41 %.
- Berdasarkan pengujian terhadap korpus yang menggunakan MADAMIRA menghasilkan nilai BLEU yang lebih besar dibandingkan sebelum

menggunakan MADAMIRA. Nilai BLEU pada korpus alquran sebesar 1,44%, korpus hadits sebesar 32,90%, dan korpus gabungan 41,46%. Penggunaan MADAMIRA pada korpus sangat berpengaruh terhadap nilai BLEU suatu korpus.

3. Pada pengujian dengan menggunakan 4 tingkat jumlah kalimat terhadap korpus yang tidak menggunakan MADAMIRA terjadi penurunan. Semakin banyak jumlah kalimat yang digunakan dalam pengujian maka semakin kecil juga nilai BLEU yang dihasilkan.
4. Berdasarkan pengujian menggunakan 4 tingkat jumlah kalimat pada korpus yang menggunakan MADAMIRA nilai BLEU yang dihasilkan pada korpus alquran terjadi kenaikan dan penurunan. Nilai BLEU pada 2000 kalimat 12,56%, 3000 kalimat menghasilkan nilai BLEU 12,38%, 4000 menghasilkan nilai BLEU 12,59 % dan 5000 kalimat menghasilkan nilai BLEU 12,33 %..
5. Pengujian 4 tingkat jumlah kalimat pada korpus hadits yang menggunakan MADAMIRA menghasilkan nilai BLEU yang kecil pada 5000 kalimat. Ketika menggunakan 5000 kalimat menghasilkan nilai BLEU sebesar 21,46 % akan tetapi pada 3000 kalimat nilai BLEU sebesar 28,47 %.
6. Pada korpus gabungan ketika dilakukan pengujian menggunakan 4 tingkat jumlah korpus nilai BLEU terhadap korpus tanpa menggunakan MADAMIRA mengalami penurunan. Semakin banyak jumlah kalimat yang digunakan pada pengujian ini semakin sedikit nilai BLEU yang dihasilkan.

IV. KESIMPULAN

Berdasarkan hasil analisis dan pengujian terhadap penggunaan Alquran dan Hadits nabi sebagai korpus mesin penerjemah, maka dapat ditarik kesimpulan sebagai berikut:

1. Korpus bahasa arab yang bersumber dari alquran dan hadits sebaiknya tidak digabungkan. Jika digabungkan hasil akurasi akan menurun.
2. Perbedaan jenis bahasa arab yang digunakan pada alquran dan hadits menjadi penyebab hasil nilai BLEU rendah.
3. Pada korpus tanpa menggunakan MADAMIRA, nilai BLEU yang dihasilkan korpus alquran sebesar 10,56 %, korpus hadits sebesar 27,65 % dan korpus gabungan sebesar 15,41%.
4. Pada korpus tanpa menggunakan MADAMIRA dilakukan pengujian 4 tingkat jumlah kalimat menunjukkan semakin banyak jumlah kalimat pada korpus maka semakin kecil nilai BLEU yang dihasilkan.
5. Hasil frasa jeda dari *shallow parsing* ini dapat membantu dalam pengembangan sistem *text to speech*. Dimana potongan-potongan frasa jeda tersebut dapat digunakan sebagai informasi untuk memproses prosodi generasi pada sistem TTS Bahasa Melayu Pontianak.

6. Pada korpus alquran yang menggunakan MADAMIRA dilakukan pengujian 4 tingkat jumlah kalimat, mengalami kenaikan nilai BLEU pada pengujian pertama dan ketiga. Pada pengujian yang kedua dan keempat mengalami penurunan nilai BLEU.
7. Pada korpus hadits yang menggunakan MADAMIRA ketika di dilakukan pengujian 4 tingkat jumlah korpus, mengalami penurunan nilai BLEU ketika jumlah kalimat sebesar 4000 dan 5000 kalimat. Nilai BLEU pada pengujian menggunakan 4000 kalimat sebesar 27,09 % dan 5000 kalimat sebesar 21,46 %.
8. Proses tokenisasi dan cleaning pada korpus mempengaruhi penilaian ahli bahasa terhadap hasil terjemahan.

REFERENSI

- [1] S. Ebrahim, D.Hegazy, M.G.M.Mostafa, and El-Beltagy S.R. (2015) English-Arabic Statistical Machine Translation: State of the Art. In: Gelbukh A. (eds) Computational Linguistics and Intelligent Text Processing. CICLing 2015. Lecture Notes in Computer Science, vol 9041. Springer, Cham.
- [2] 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.
- [3] Y. Jarob, H. Sujaini dan N. Safriadi, "Uji Akurasi Penerjemahan Bahasa Indonesia – Dayak Taman dengan Penandaan Kata Dasar dan Imbuan", *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, Vol. 2, No. 2, 2016.
- [4] 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.
- [5] M, Christopher D. dan Schutze, Hinrich. 2000. Foundations Of Statistical Natural Language Processing. London : The MIT Press Cambridge Massachusetts.
- [6] H. Sujaini, A,B, Putra. 2015. *Analysis of Extended Word Similarity Clustering based Algorithm on Cognate Language*. Gujarat: ESRSA Publications Pvt. Ltd.
- [7] S. Hunston. 2002. Corpora in Applide Linguistics. Cambridge: Cambrigde University Press.
- [8] K. Papineni, dkk. 2002. BLEU: *a Method for Automatic Evaluation of Machine Translation*. USA: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, July 2002, pp. 311-318.
- [9] H. Tanuwijaya. 2009. *Penerjemahan Inggris-Indonesia Menggunakan Mesin Penerjemah Statistik Dengan Word Reordering dan Phrase Reordering*. Jakarta : Universitas Indonesia.