

IMPLEMENTASI MESIN PENERJEMAH STATISTIK BERBASIS ANDROID DENGAN MOSES DECODER

Try Wahyudinata¹, Herry Sujaini², Rudy Dwi Nyoto³
 Program Studi Teknik Informatika Universitas Tanjungpura^{1,2,3}
e-mail: trywahyudinata@gmail.com¹, herry_sujaini@yahoo.com², rudydn@gmail.com³

Abstrak - Indonesia memiliki keragaman bahasa yang luar biasa. Akan tetapi, keterbatasan dalam menguasai berbagai macam bahasa daerah di Indonesia merupakan kendala utama bagi masyarakat umum untuk bisa mengerti bahasa daerah yang bukan berasal dari daerah asalnya. Mesin penerjemah statistik (*Statistical Machine Translation*) merupakan sebuah pendekatan mesin penerjemah dengan hasil terjemahan yang dihasilkan atas dasar model statistik yang parameter-parameternya diambil dari hasil analisis korpus teks paralel. Saat ini telah banyak dilakukan penelitian mengenai penerjemahan bahasa-bahasa daerah ke bahasa Indonesia. Dalam penelitian-penelitian tersebut hanya menguji nilai akurasi hasil terjemahan dari mesin yang dibangun dan belum dapat dipublikasikan dan diakses oleh publik. Tujuan penelitian ini adalah mengimplementasikan mesin penerjemah statistik untuk menerjemahkan bahasa secara satu arah maupun dua arah berbasis Android dengan menggunakan Moses Decoder. Sehingga dapat diakses secara luas oleh publik kapan saja, di mana saja hanya dengan menggunakan *smartphone* Android yang terkoneksi dengan internet dalam sebuah aplikasi penerjemah bahasa. Berdasarkan hasil implementasi, aplikasi ini sudah berjalan dengan baik pada perangkat *mobile* Android, mulai dari Android versi 4.0.3 (*Ice Cream Sandwich*) hingga Android versi 6.0.1 (*Marsmallow*). Berdasarkan hasil kuesioner menggunakan skala Likert's Summated Rating (LSR) diperoleh nilai total 1477 yang berarti aplikasi ini diimplementasikan dengan sangat baik.

Kata Kunci : Mesin Penerjemah Statistik, Android, Moses Decoder, Aplikasi Penerjemah Bahasa.

I. PENDAHULUAN

Indonesia merupakan negara yang memiliki keragaman budaya dan bahasa. Bahasa Indonesia merupakan bahasa nasional yang mempersatikan seluruh masyarakat Indonesia. Faktor identitas sosial masyarakat seperti perbedaan status, kedudukan ekonomi, pendidikan dan lain-lain menyebabkan adanya keragaman budaya dan bahasa. Sebagai makhluk sosial manusia perlu berinteraksi dengan manusia lainnya menggunakan bahasa untuk menyampaikan apa yang mereka maksud. Bahasa yang digunakan bertujuan untuk menyampaikan ide, keinginan, perasaan maupun pengalaman kepada orang lain.

Dalam berkomunikasi dan berinteraksi, manusia menggunakan bahasa yang dapat dimengerti satu sama lain. Indonesia memiliki keberagaman bahasa dan budaya yang sangat luar biasa. Ada sebanyak 726 bahasa daerah dengan 719 bahasa daerah diantaranya yang masih aktif digunakan sehari-hari diseluruh pelosok Indonesia [1]. Tetapi kemampuan menguasai berbagai macam bahasa yang terbatas merupakan masalah yang umum. Penerjemahan suatu bahasa ke bahasa lain dengan penerjemahan per kata dianggap tidak selalu pas atau

akurat untuk menerjemahkan bahasa tersebut. Tidak jarang akan menghasilkan penafsiran yang aneh dan membuat hilangnya arti dan maksud yang terkandung dari sebuah kalimat. Saat ini telah banyak dilakukan penelitian mengenai penerjemahan bahasa-bahasa daerah ke bahasa Indonesia. Dalam penelitian-penelitian tersebut hanya menguji nilai akurasi hasil terjemahan dari mesin yang dibangun dan belum dapat dipublikasikan dan diakses oleh publik.

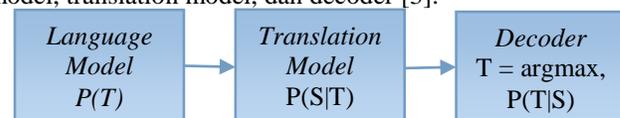
II. URAIAN PENELITIAN

A. Penerjemahan

Dalam Kamus Besar Bahasa Indonesia (KBBI) kata “terjemah/ menerjemahkan” merupakan menyalin (memindahkan) suatu bahasa ke bahasa lain atau mengalihbahasakan. Penerjemahan adalah kegiatan mengalihkan secara tertulis pesan dari teks suatu bahasa (misalnya bahasa Inggris) ke dalam teks bahasa lain (misalnya bahasa Indonesia) [2]. Berdasarkan uraian di atas, dapat disimpulkan bahwa penerjemahan adalah mengubah atau mengalihkan suatu teks atau pesan dari bahasa sumber ke bahasa target yang berbeda tanpa mengubah makna dari teks atau pesan tersebut.

B. Mesin Penerjemah

Mesin penerjemah (*machine translation*) merupakan alat penerjemah otomatis pada sebuah teks dari satu bahasa ke bahasa lainnya. Ada beberapa pendekatan untuk *machine translation* seperti pendekatan dengan menggunakan aturan (*rule-based machine translation*), pendekatan dengan menggunakan contoh (*example-based machine translation*), dan pendekatan dengan menggunakan model statistik (*statistical machine translation*). Dalam mesin penerjemah statistik, terdapat 3 komponen yang terlibat dalam proses penerjemahan dari satu bahasa ke bahasa lain yaitu : *language model*, *translation model*, dan *decoder* [3].



Gambar 1. Komponen Mesin Penerjemah Statistik

C. Moses Decoder

Moses merupakan implementasi dari pendekatan dengan menggunakan metode statistik (*statistical machine translation*). Pendekatan ini adalah pendekatan yang sering dipakai pada saat ini, dan telah dipakai oleh sistem penerjemahan oleh Google dan

Microsoft. Pada statistical machine translation (SMT), sistem penerjemahan melakukan training pada sekumpulan korpus paralel. Korpus paralel adalah kumpulan dari kalimat-kalimat dalam dua bahasa yang berbeda, yang merupakan kalimat-kalimat selaras, setiap kalimat dalam satu bahasa cocok dengan kalimat yang diterjemahkan dalam bahasa lain. Moses Decoder bekerja untuk menemukan skor tertinggi dari kalimat pada bahasa target berdasarkan model translasi yang sesuai dengan bahasa sumber. Model bahasa dibangun dengan korpus monolingual bahasa target menggunakan SRILM, KenLM ataupun IRSTLM. Selain itu, dengan menggunakan GIZA++ dibangun model translasi dan susunan kata untuk penerjemahan yang di-training dari korpus paralel [4].

D. Unified Modelling Language (UML)

Unified Modeling Language (UML) adalah standard pemodelan dalam pembuatan aplikasi berorientasi objek yang diajukan oleh *Object Management Group* (OMG) pada tahun 1996 [5].

1. Use Case Diagram

Use case diagram mendeskripsikan sebuah interaksi antara satu atau lebih aktor dengan sistem informasi yang akan dibuat. Dengan kata lain, use case diagram digunakan untuk mengetahui fungsi-fungsi apa saja yang terdapat di dalam sistem dan siapa saja yang berhak mengakses fungsi tersebut [5].

2. Class Diagram

Diagram kelas atau class diagram menggambarkan struktur sistem dari segi pendefinisian kelas-kelas yang akan dibuat untuk membangun sistem. Kelas memiliki apa yang disebut atribut dan metode atau operasi. Atribut merupakan variabel-variabel yang dimiliki oleh suatu kelas. Metode atau operasi adalah fungsi-fungsi yang dimiliki oleh suatu kelas [5].

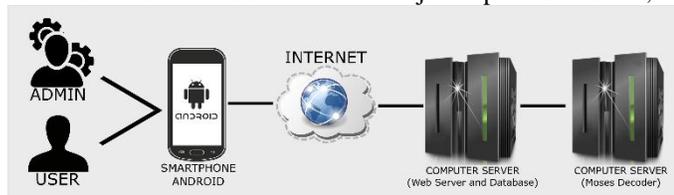
3. Sequence Diagram

Sequence diagram menggambarkan kelakuan objek pada use case dengan mendeskripsikan waktu hidup objek dan message yang dikirimkan dan diterima antar objek. Dalam menggambarkan sequence diagram perlu memperhatikan objek-objek yang terlibat di dalam use case beserta metode-metode yang dimiliki kelas yang diinstansiasi menjadi objek itu [5].

III. PERANCANGAN SISTEM

A. Arsitektur Sistem

Desain arsitektur sistem akan ditunjukkan pada Gambar 2,



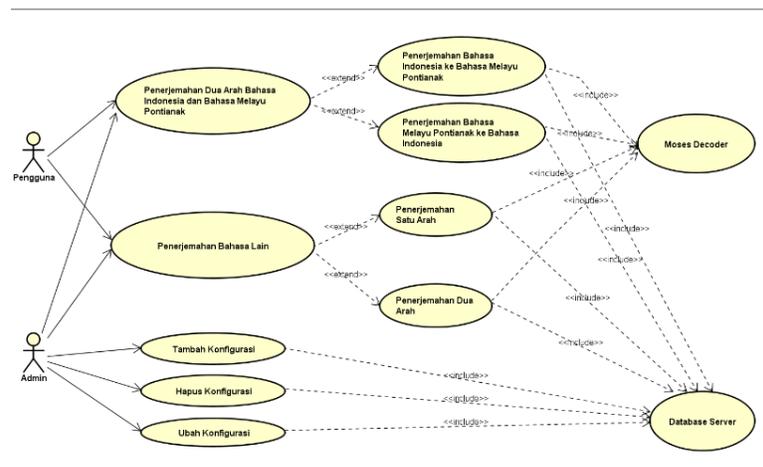
Gambar 2. Desain Arsitektur Sistem

Melalui internet, aplikasi android pengguna (Admin dan User) akan melakukan komunikasi data dengan web server. Pengguna mengirimkan masukan berupa kata/kalimat ke web server melalui internet dan diproses di Moses Decoder yang

menghasilkan terjemahan dan dikirim kembali ke aplikasi android.

B. Use Case Diagram

Use case diagram aplikasi diperlihatkan pada Gambar 3,



Gambar 3. Use Case Diagram Aplikasi

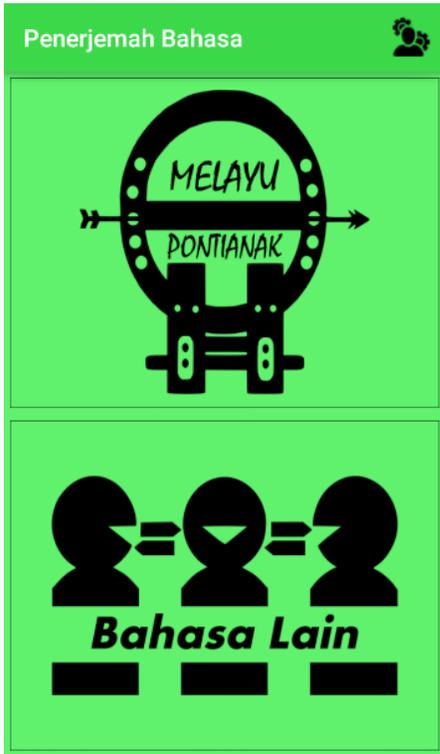
C. Pengujian Aplikasi

Pengujian aplikasi dilakukan dengan menggunakan metode Black Box, kompatibilitas aplikasi dan kuesioner terhadap 25 responden yang terdiri dari mahasiswa.

D. Hasil Aplikasi

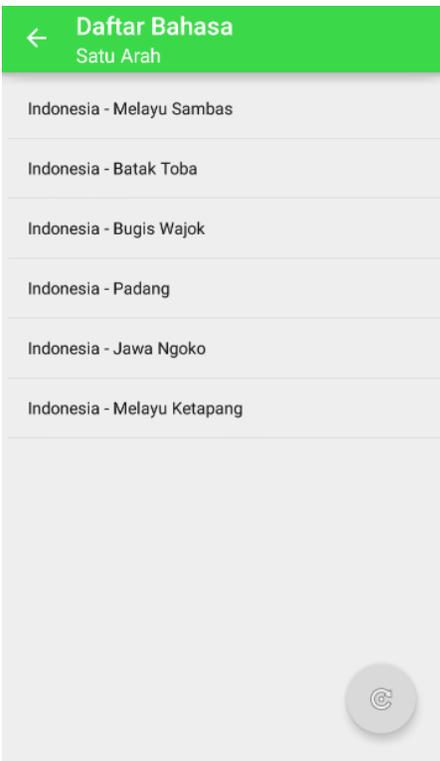
Aplikasi yang dibangun merupakan implementasi dari mesin penerjemah statistik Moses Decoder dengan basis Android. Dalam implementasinya, mesin penerjemah statistik dengan bahasa yang sudah di bangun atau di-training memerlukan suatu masukan (*input*) yang berupa kata / kalimat yang selanjutnya akan diproses oleh mesin penerjemah statistik Moses Decoder yang telah di-training dengan Srilm dan Giza++. Kemudian akan menghasilkan suatu keluaran (*output*) berupa hasil terjemahan kata / kalimat dari masukan (*input*) sebelumnya.

Aplikasi ini dibuat agar dapat diimplementasikan dengan banyak bahasa (*multi-language*). Sehingga dapat ditambahkan bahasa selain bahasa yang sudah tersedia secara *default* pada aplikasi ini oleh admin. Bahasa yang di tambahkan dapat dikonfigurasi dengan penerjemahan satu arah dan penerjemahan dua arah. Untuk penerjemahan bahasa yang tetap, digunakan penerjemahan bahasa dua arah bahasa Indonesia dan bahasa Melayu Pontianak. Berikut beberapa tampilan hasil perancangan aplikasi, yang diperlihatkan pada Gambar 3 hingga Gambar 7.



Gambar 4. Tampilan Menu Utama

Pada Gambar 4, merupakan tampilan menu utama pada aplikasi ini. Menu utama memiliki 3 menu yaitu, menu login admin, menu penerjemahan bahasa Melayu Pontianak dan menu penerjemahan bahasa lain (Secara satu arah dan dua arah).



Gambar 5. Tampilan Daftar Bahasa

Gambar 5 merupakan tampilan daftar bahasa dari penerjemahan satu arah.



Gambar 6. Tampilan Tambah Konfigurasi

Gambar 6 merupakan tampilan untuk tambah konfigurasi bahasa oleh admin.



Gambar 7. Tampilan Penerjemahan Satu Arah

Gambar 7 merupakan tampilan penerjemahan bahasa satu arah dari bahasa Indonesia ke bahasa Melayu Sambas.



Gambar 8. Tampilan Penerjemahan Dua Arah

Gambar 8 merupakan tampilan penerjemahan dua arah, bahasa Indonesia – bahasa Melayu Sambas.

E. Hasil Pengujian

1. Robustness Testing

Robustness Testing adalah pengujian dengan data *input* dipilih diluar spesifikasi yang telah didefinisikan. Tujuan dari pengujian ini adalah membuktikan bahwa tidak ada kesalahan jika masukan tidak valid. Pengujian ini dilakukan pada proses input data seperti penerjemahan bahasa dan tambah konfigurasi bahasa [5]. Pada Tabel 1 memperlihatkan pengujian pada penerjemahan bahasa Melayu Pontianak ke bahasa Indonesia dalam pengujian dilakukan dengan data kosong, satu kata, dua kata dan satu kalimat.

Tabel 1
Tabel Pengujian Penerjemahan Bahasa

Input	Contoh Data	Hasil Eksekusi	Keterangan
Data kosong		Tidak berhasil	Pesan kesalahan: Masukkan teks yang akan diterjemahkan!
Satu kata	Aku	Berhasil. Terjemahan: Aku	
Dua kata	Aku dah	Berhasil. Terjemahan: Aku sudah	
Satu kalimat	Aku udah suah pegi ke tugu katulistiwa	Berhasil. Terjemahan: Aku sudah pernah pergi ke tugu khatulistiwa	

Pada Tabel 2, memperlihatkan pengujian tambah konfigurasi bahasa. Dalam pengujian ini dilakukan dengan data kosong, sebagian data kosong dan data lengkap.

Tabel 2
Tabel Hasil Pengujian Tambah Konfigurasi

Input	Contoh Data		Hasil Eksekusi	Ket
Data kosong			Tidak berhasil	Pesan kesalahan: Isi kolom Bahasa Masukan!
Sebagian data kosong	Bahasa masukan	Indonesia	Tidak berhasil	Pesan kesalahan: Isi kolom Bahasa Keluaran!
	Bahasa keluaran			
	IP Server	203.24.50.141		
	Port Server	1234		
Data lengkap	Bahasa masukan	Indonesia	Berhasil	Pesan: Berhasil!
	Bahasa keluaran	Sunda		
	IP Server	203.24.50.141		
	Port Server	2109		

2. Kompatibilitas Aplikasi

Pengujian ini dilakukan dengan menginstal aplikasi pada beberapa perangkat android yang berbeda.

Tabel 3
Pengujian Kompatibilitas Aplikasi

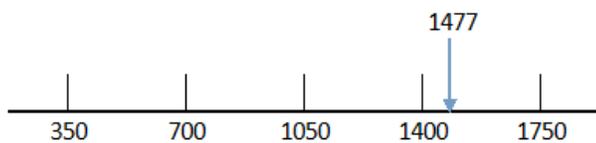
No	Merek Perangkat	Layar	Versi Android	Ket.
1	HTC Sensation XE Z715e	Layar 4.3 inches, resolusi: 540 x 960 pixels	4.0.3 (Ice Cream Sandwich)	Aplikasi Berjalan Lancar
2	Axioo Picopad GFI	Layar 6.0 inches, resolusi: 854x480 pixels	4.1.1 (Jelly Bean)	Aplikasi Berjalan Lancar
3	Samsung GT-S6310 (Galaxy Young)	Layar 3.27 inches, resolusi: 320x480 pixels	4.1.2 (Jelly Bean)	Aplikasi Berjalan Lancar
4	Smartfren Andromax C	Layar 4.0 inches, resolusi: 480x800 pixels	4.1.2 (Jelly Bean)	Aplikasi Berjalan Lancar
5	ASUS T00G (Zenfone 6)	Layar 6.0 inches, resolusi: 720x1280 pixels	4.4.2 (KitKat)	Aplikasi Berjalan Lancar
6	Samsung SM-T311 (Galaxy Tab 3 8.0)	Layar 8.0 inches, resolusi: 800x1280 pixels	4.4.2 (KitKat)	Aplikasi Berjalan Lancar
7	Advance S50	Layar 5.0 inches, resolusi: 854x480 pixels	4.4.2 (KitKat)	Aplikasi Berjalan Lancar
8	ASUS Z007 (Zenfone 4C)	Layar 4.5 inches, resolusi: 480x854 pixels	4.4.2 (KitKat)	Aplikasi Berjalan Lancar
9	ASUS T00I (Zenfone 4)	Layar 4.0 inches, resolusi: 480x800 pixels	4.4.2 (KitKat)	Aplikasi Berjalan Lancar
10	V-Gen Birdie A2	Layar 4.0 inches, resolusi: 480x800 pixels	4.4.2 (Kitkat)	Aplikasi Berjalan Lancar
11	Samsung SM-T231 (Galaxy Tab 4 7.0)	Layar 7.0 inches, resolusi: 800x1280 pixels	4.4.2 (KitKat)	Aplikasi Berjalan Lancar
12	Sony E1 D2005	Layar 4 inches, resolusi: 480x800 pixels	4.4.2 (KitKat)	Aplikasi Berjalan Lancar
13	Xiaomi Mi4	Layar 5.0 inches, resolusi: 1080x1920 pixels	4.4.4 (KitKat)	Aplikasi Berjalan Lancar
14	Xiaomi Redmi 2 Prime	Layar 4.7 inches, resolusi: 720x1280 pixels	4.4.4 (KitKat)	Aplikasi Berjalan Lancar
15	ASUS Z008D (Zenfone 2)	Layar 5.5 inches, resolusi: 720x1280 pixels	5.0.0 (Lollipop)	Aplikasi Berjalan Lancar
16	Xiaomi Redmi Note 3	Layar 5.5 inches, resolusi: 1080x1920 pixels	5.0.2 (Lollipop)	Aplikasi Berjalan Lancar
17	Lenovo A6000+	Layar 5.0 inches, resolusi: 720x1280 pixels	5.0.2 (Lollipop)	Aplikasi Berjalan Lancar
18	Xiaomi Mi4i	Layar 5.0 inches, resolusi: 1080x1920 pixels	5.0.2 (Lollipop)	Aplikasi Berjalan Lancar
19	Samsung J7	Layar 5.5 inches, resolusi: 720x1280 pixels	5.1.2 (Lollipop)	Aplikasi Berjalan Lancar
20	Xiaomi Mi4	Layar 5.0 inches, resolusi: 1080x1920 pixels	6.0.1 (Marshmallow)	Aplikasi Berjalan Lancar

3. Kuesioner

Kuesioner merupakan teknik pengumpulan data yang dilakukan dengan cara memberi seperangkat pertanyaan atau pernyataan tertulis kepada responden untuk dijawab [6]. Berikut adalah total skor dari kuesioner yang telah dibagikan kepada 25 responden.

Tabel 4
Total Skor Responden Dari Kuesioner

Responde n	Item														Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
A	5	5	5	5	4	5	5	5	5	5	4	5	4	5	67
B	4	5	5	4	4	5	5	5	4	4	4	4	5	4	63
C	5	5	5	5	4	5	4	4	4	4	4	4	4	4	61
D	5	5	5	3	3	4	5	5	5	4	4	5	4	5	62
E	5	5	5	4	4	4	5	5	4	4	4	5	4	5	63
F	5	5	5	5	5	5	5	5	4	4	4	5	4	5	67
G	5	5	5	5	5	5	5	5	4	4	5	5	5	5	68
H	4	4	4	4	4	4	3	3	4	3	3	3	2	3	48
I	5	5	5	5	4	4	4	4	4	4	4	4	3	3	59
J	5	5	5	5	5	5	5	5	5	4	4	5	4	5	67
K	5	4	5	4	4	5	5	4	5	3	4	4	4	5	61
L	4	5	4	5	4	5	5	5	4	4	4	5	4	5	63
M	3	5	5	4	4	5	5	4	3	4	4	4	3	4	57
N	5	5	5	5	5	5	5	5	5	2	2	4	2	5	60
O	4	4	4	4	4	4	4	4	3	4	3	4	3	5	54
P	5	5	5	5	4	5	5	5	3	4	3	5	3	5	62
Q	3	3	3	4	4	4	4	4	4	3	3	4	4	5	52
R	4	5	5	4	4	4	5	5	4	3	4	3	3	4	57
S	5	5	5	5	5	5	4	4	4	4	4	5	4	4	63
T	3	4	4	4	5	4	4	4	4	3	3	4	3	4	53
U	5	5	5	4	3	4	4	4	4	3	3	5	4	5	58
V	4	5	4	4	4	4	4	4	4	4	5	4	4	4	58
W	4	5	4	4	4	4	4	5	4	4	5	4	3	5	59
X	5	4	4	3	4	3	4	4	4	3	3	4	2	4	51
Y	4	4	4	3	3	4	3	3	3	2	3	2	3	4	44
Total Skor															1477



Gambar 9. Hasil Kuesioner Pada Interpretasi LSR

F. Analisis Hasil Pengujian

Rincian hasil analisis pengujian aplikasi penerjemah bahasa berbasis Android yang telah dilakukan adalah sebagai berikut:

1. Hasil *robustness testing* menyatakan aplikasi dapat meng-handle proses *input* dengan baik, terutama pada proses *input* data yang tidak sesuai dan kosong.
2. Berdasarkan hasil pengujian kompatibilitas aplikasi, diperoleh bahwa aplikasi dapat berjalan pada perangkat *smartphone* Android dengan sistem operasi versi 4.0.3 (Ice Cream Sandwich) hingga versi 6.0.1 (Marsmallow).
3. Berdasarkan hasil pengujian kompatibilitas aplikasi, diperoleh hasil bahwa tampilan pada setiap perangkat *smartphone* Android memiliki perbedaan, hal ini disebabkan perbedaan ukuran dan resolusi pada setiap layar *smartphone* Android.
4. Berdasarkan hasil pengujian kompatibilitas aplikasi, sistem tidak berjalan dengan baik pada saat menampilkan daftar bahasa yang ada di server. Kurang stabilnya koneksi internet yang ada menyebabkan hal ini terjadi.
5. Hasil kuesioner menunjukkan sebagian besar responden menerima dengan baik aplikasi ini dari segi pengoperasian, fungsionalitas dan tampilan.
6. Responden memberikan beberapa masukan untuk perbaikan aplikasi seperti tambahan fitur aplikasi dan perbaikan pada tampilan aplikasi.

7. Hasil interpretasi skor dari *Likert's Summated Rating* (LSR) menunjukkan hasil 1477 dengan nilai sangat positif yang berarti aplikasi dinilai berhasil.
8. Hasil uji validitas pada kuesioner menunjukkan bahwa setiap *item* pada kuesioner bernilai valid karena nilai korelasi setiap *item* (r_{hitung}) pada kuesioner lebih besar dari nilai korelasi tabel (r_{tabel}) untuk 25 responden.

Hasil uji reliabilitas pada kuesioner menunjukkan bahwa kuesioner memiliki reliabilitas yang signifikan atau dengan kata lain reliabilitas kuesioner baik, yang berarti data hasil kuesioner dapat dipercaya dengan nilai reliabilitas sebesar 0,893 yaitu lebih besar dari nilai korelasi tabel (r_{tabel}) untuk 25 responden dengan taraf signifikan 5%.

IV. KESIMPULAN/RINGKASAN

Berdasarkan hasil implementasi dan hasil analisis pengujian terhadap aplikasi penerjemah bahasa berbasis android dapat disimpulkan bahwa:

1. Mesin Penerjemah Statistik dapat diimplementasikan untuk menterjemahkan bahasa berbasis Android dengan menggunakan Moses Decoder.
2. Berdasarkan hasil pengujian kompatibilitas, aplikasi dapat berjalan lancar pada berbagai jenis versi Android yaitu versi 4.0.3 (Ice Cream Sandwich) hingga versi 6.0.1 (Marsmallow).

Berdasarkan hasil kuesioner menggunakan skala *Likert's Summated Rating* (LSR) diperoleh nilai total 1477 yang berarti aplikasi ini diimplementasikan dengan sangat baik.

DAFTAR PUSTAKA

- [1] Akuntono, Indra. 2012. *Mau Tahu Jumlah Ragam Bahasa di Indonesia?*. Kompas.com. 20 Januari 2016.
- [2] Amalia, Farida. 2007. *Ideologi Dalam Penerjemahan*. Universitas Pendidikan Indonesia. 20 Januari 2016.
- [3] Manning, Christopher D., Schutze, Hinrich. 2000. *Foundations Of Statistical Natural Language Processing*. London : The MIT Press Cambridge Massachusetts
- [4] Moses Decoder. _____. Moses – Moses Overview. <http://www.statmt.org/moses/?n=Moses.Overview> (diakses tanggal 15 Mei 2016).
- [5] Sukamto, Rosa Ariani dan M. Shalahudin. 2013. *Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek*. Bandung: Informatika
- [6] Sugiyono. 2003. *Metode Penelitian Bisnis*. Bandung. Pusat Bahasa Depdiknas.