

PENYUSUNAN *STRONG'S CONCORDANCE* UNTUK ALKITAB PERJANJIAN BARU BAHASA INDONESIA

Gunawan¹, Devi Dwi Purwanto², Herman Budianto², dan Indra Maryati²

¹Jurusan Teknik Elektro, Fakultas Teknologi Industri, Institut Teknologi Sepuluh Nopember, Surabaya 60111
admin@hansmichael.com

²Jurusan Teknik Informatika, Sekolah Tinggi Teknik Surabaya, Surabaya 60284
devi@stts.edu, herman@stts.edu, indra.maryati@gmail.com

Abstrak

Strong's Concordance adalah konkordansi yang digunakan sebagai alat bantu untuk menunjuk kata tersebut ke dalam bahasa aslinya sehingga dengan demikian dapat diketahui makna dari firman Tuhan yang hendak disampaikan. Pada penelitian ini dilakukan penyusunan *Strong's Concordance* tersebut ke dalam Alkitab Perjanjian Baru Bahasa Indonesia yang sampai saat ini belum pernah ditemukan Alkitab Perjanjian Baru Bahasa Indonesia secara *online* yang dilengkapi dengan *Strong's Concordance*. Penyusunan *Strong's Concordance* tersebut dilakukan dengan menggunakan pedoman teori yang ada pada *Natural Language Processing* (NLP) dan teori *Web Mining*. Penyusunan nomor *strong* tersebut dimulai dengan melakukan pendekatan nomor *strong* berdasarkan kemunculan katanya. Kemudian pada tahap selanjutnya digunakan pendekatan *alignment* antara kata yang ada pada Alkitab Bahasa Indonesia dengan nomor *strong* yang terdapat pada Alkitab Bahasa Inggris dengan menggunakan *word alignment*. Pendekatan ketiga menggunakan pendekatan *n-gram* dengan perhitungan *mutual information* untuk mencari arti kata yang terdiri lebih dari satu kata. Pendekatan keempat dilakukan dengan cara melakukan *stemming* pada *corpus* Alkitab Perjanjian Baru Bahasa Indonesia yang mana nantinya digunakan sebagai *corpus* baru untuk melakukan pencarian pada tahap satu sampai dengan tahap tiga. Maksud dari *stemming* tersebut adalah untuk mendapatkan *root* dari arti kata tersebut. Selain keempat pendekatan tersebut juga dilakukan pendekatan lain seperti pencarian *proper name*, pencarian nomor *strong* yang hanya memiliki satu frekuensi dan pendataan nomor *strong* yang termasuk dalam *conjunction*, *preposition*, dan *pronoun*.

Kata kunci: *Concordance*, *Natural Language Processing*, *N-Gram*, *Vector Space Model*.

1. Pendahuluan

Alkitab adalah kumpulan firman Tuhan yang digunakan sebagai panduan hidup umat Kristen. Alkitab juga merupakan buku yang paling banyak diterjemahkan, yaitu ke dalam 2179 dialek dan bahasa [1]. Dari hasil penterjemahan ke dalam berbagai bahasa tersebut, didapati penterjemahan yang kurang tepat. Hal ini disebabkan karena perbedaan *vocabulary* atau perbedaan antara bahasa yang satu dengan yang lain. Dari sini munculnya ide menggunakan nomor *strong*.

Dengan menggunakan konkordansi, kita dapat melihat di mana dan bagaimana kata tertentu digunakan. Dengan demikian, kita dapat melihat arti dari sebuah kata dengan tepat. Sebagai contoh adalah kata *agape* (kasih). Orang tertentu mengartikan kata

tersebut sebagai kasih Allah yang mulia dan suci. Hal itu benar, tetapi tidak selamanya demikian. Karena Paulus menulis kepada Timotius: "Karena Demas telah mencintai dunia ini dan meninggalkan aku ..." (IITim. 4:10). Kata *mencintai*, dalam bahasa aslinya pada ayat tersebut di atas adalah *agape*. Apakah ini berarti bahwa Demas mencintai dunia ini dengan kasih *agape*? Jawabnya tentu tidak. Oleh karena itu pemahaman tentang arti kata tersebut sangatlah penting, agar tidak salah memahami pesan yang tersirat dalam Alkitab.

Pada penelitian ini dilakukan pembuatan aplikasi *web* yang dapat membantu seseorang dalam memahami Alkitab. Aplikasi *web* ini berfungsi untuk mencari kesamaan ayat dan mengetahui kata asli dari Bahasa Yunani dari nomor *strong* sebelum diterjemahkan.

2. Strong's Concordance

2.1. Pengertian Concordance

Konkordansi adalah daftar menurut abjad dari kata utama yang dipakai pada buku/dokumentasi dengan konteks yang telah ditentukan. Karena kesulitan waktu dan biaya yang dibutuhkan dalam menciptakan sebuah konkordansi di era pra-komputer, hanya karya-karya penting khusus, seperti Alkitab, Al Qur'an, atau karya-karya Shakespeare yang telah dilakukan konkordansi. Walaupun penomoran *strong* dilakukan secara terkomputerisasi, hasil dari penomoran konkordansi tersebut diperlukan banyak pekerjaan manual [2].

Strong's Concordance atau *Strong's Exhaustive Konkordansi Alkitab* adalah konkordansi Alkitab yang dibangun di bawah arahan Dr James Strong (1822-1894) dan pertama kali diterbitkan pada tahun 1890. Tujuan dari konkordansi ini tidak menyediakan konten atau keterangan dari Alkitab, tetapi untuk memberikan indeks Alkitab. Hal ini memungkinkan pembaca untuk menemukan kata-kata untuk melihat ayat dalam Alkitab. Indeks ini juga memungkinkan pembaca untuk mencari sebuah frasa untuk membandingkan bagaimana topik yang sama dibahas dalam bagian yang berbeda dari Alkitab.

Penggunaan nomor *Strong* tidak mempertimbangkan kiasan, metafora, idiom, frasa umum, budaya referensi, referensi peristiwa sejarah, atau makna-makna alternatif yang digunakan oleh orang-orang dalam jangka waktu tertentu untuk mengungkapkan pikiran mereka dalam bahasa mereka sendiri pada saat itu.

2.2. Keuntungan Concordance

Konkordansi biasanya digunakan sebagai alat dalam linguistik yang digunakan dalam teks atau korpus belajar yang didokumentasikan dengan baik. Didokumentasikan dengan baik berarti memiliki struktur yang jelas, yang dapat digunakan untuk menghubungkan dua bahasa yang berbeda korpus.

Ada empat keuntungan dari Konkordansi:

1. Membandingkan penggunaan kata-kata yang berbeda dari kata dasar yang sama.
2. Meneliti kata kunci.
3. Meneliti frekuensi kata-kata yang digunakan dalam dokumentasi.
4. Membuat indeks dan daftar kata yang digunakan dalam Alkitab.

2.3. Keuntungan Strong's Concordance

Strong's Concordance digunakan untuk mencari nomor *Strong* sehingga dapat digunakan untuk memberikan indeks Alkitab. Ada empat keuntungan dari *Strong's Concordance*:

1. Menemukan dan meneliti frase dan idiom.
2. Menemukan terjemahan *subsential* E.G. terminologi dalam dua bahasa yang berbeda korpus.
3. Membantu untuk mengetahui kebenaran Alkitab dalam bahasa aslinya sehingga tidak ada yang salah paham apa firman Allah.
4. Membantu orang dalam menafsirkan Alkitab.

3. Natural Language Processing

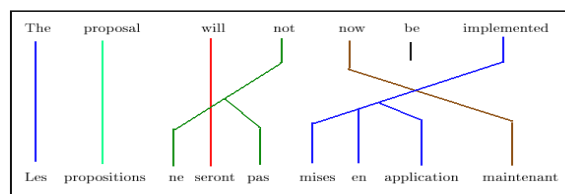
Natural Language Processing (NLP) adalah proses transformasi informasi yang diekspresikan dalam bahasa pembicaraan dan tertulis dari masyarakat untuk dimasukkan ke komputer melalui perangkat lunak untuk memperoleh informasi tertentu. Dalam penelitian ini akan menjelaskan mengenai *statistical NLP* (statistik NLP) dan *text retrieval* (pengambilan teks). Keduanya akan jelaskan di bawah.

3.1. Statistical NLP

Statistik NLP melakukan analisis statistik mengenai korpus yang ada, dengan asumsi bahwa sebagian besar naskah-naskah (korpus) yang bisa menggambarkan sifat bahasa yang digunakan dalam informasi statistik *corpus*. Ada 3 metode statistik yang digunakan untuk memperoleh informasi statistik mengenai korpus.

Word Alignment

Word alignment adalah salah satu *task* dalam NLP yang digunakan untuk mengidentifikasi hubungan translasi atau terjemahan antara pasangan kata dalam dua korpus yang berbeda bahasa. *Word alignment* biasanya dilakukan setelah *sentence alignment* selesai dilakukan. *Sentence alignment* sendiri adalah proses untuk mengidentifikasi pasangan kalimat yang mana sebuah kalimat adalah hasil translasi dari kalimat lainnya. Contoh *word alignment* dapat dilihat pada Gambar 1.



Gambar 1. Contoh *Word Alignment* [3]

Contoh tersebut menggambarkan kutipan hasil translasi antara bahasa Inggris dengan bahasa Perancis yang tidak setiap kata pada bahasa Perancis memiliki arti, demikian sebaliknya dengan Bahasa Inggris yang tidak selalu dapat ditranslasikan ke dalam Bahasa Perancis, dan tidak selalu pada kata

pertama dalam Bahasa Inggris memiliki arti pada kata pertama pada Bahasa Perancis.

Untuk mengidentifikasi hubungan translasi antara dua bahasa tersebut digunakan pendekatan *chi-square* pada matriks 2x2. Rumus pendekatan tersebut adalah:

$$X^2 = \frac{N(O_{11}O_{22} - O_{12}O_{21})^2}{(O_{11} + O_{12})(O_{11} + O_{21})(O_{12} + O_{22})(O_{21} + O_{22})} \quad (1)$$

| | | |
|---------|----------|----------|
| | word | ¬word |
| Number | O_{11} | O_{12} |
| ¬Number | O_{21} | O_{22} |

Gambar 2. Matriks 2x2

Pada kasus ini n adalah jumlah ayat pada Alkitab. Sedangkan nilai O_{11} , O_{12} , O_{21} , O_{22} dapat dilihat pada matriks 2x2 pada Gambar 2. Nilai O_{11} tersebut didapatkan dengan cara mencari jumlah frekuensi yang pada ayat tersebut terdapat pasangan kata dan nomor *strong* yang terdapat pada Alkitab Perjanjian Baru Bahasa Inggris dan kata terdapat pada Alkitab Perjanjian Baru Bahasa Indonesia. Sedangkan O_{12} didapatkan dengan mencari jumlah ayat yang memiliki nomor yang sedang dicari pada Alkitab Bahasa Inggris tetapi tidak memiliki kata pada ayat Alkitab Bahasa Indonesia. O_{21} didapatkan dengan mencari jumlah ayat yang memiliki kata pada Alkitab Bahasa Indonesia tetapi tidak memiliki nomor yang sedang dicari pada ayat Alkitab Bahasa Inggris. O_{22} didapatkan dengan mencari jumlah ayat yang tidak ada hubungannya dengan pasangan ini, yaitu yang tidak memiliki nomor pada ayat Alkitab Bahasa Inggris dan juga tidak memiliki kata pada ayat Alkitab bahasa Indonesia. Dari nilai X_2 yang didapat, nilai yang paling besar yang dicurigai sebagai kandidat kata dalam proses *alignment* tersebut. Hal tersebut dikarenakan semakin besar nilai *alignment* tersebut, maka semakin besar pula keterkaitan kata tersebut dengan nomor *strong*-nya.

N-Gram

N-Gram adalah sebuah operasi model statik yang berhubungan dengan pemodelan bahasa yang meliputi pemberian nilai pada kata atau kalimat yang pemodelan tersebut didasarkan pada jenis aplikasi yang akan dibuat. Salah satu teori untuk menyusun pemodelan bahasa tersebut adalah *Markov Models*, yang mana penyusunan pemodelan bahasa tersebut didasarkan pada teori n-gram ini.

Tujuan dari penggunaan n-gram dalam penelitian ini adalah dengan mempelajari *corpus* yang tersedia kita dapat mengetahui kandidat dari tiap nomor

strong terhadap kemunculan pasangan katanya, hal ini dikarenakan terdapat berbagai pasangan kata yang membentuk arti kata yang baru. Pengelompokkan pasangan tersebut yang umum digunakan untuk $n=2$ yang disebut dengan "bigram" dan $n=3$ yang disebut dengan "trigram".

Mutual Information

Dalam penelitian ini pengelompokkan n-gram berdasarkan huruf tidak dapat digunakan sehingga hanya pengelompokkan kata yang dapat dipakai. Kandidat tersebut nantinya akan dilakukan perhitungan dengan menggunakan bantuan *mutual information* untuk mengetahui kandidat mana yang paling memungkinkan untuk penomoran *strong* tersebut.

3.2. Text Retrieval

Text retrieval adalah suatu metode yang digunakan untuk membantu user dalam mencari suatu informasi yang berguna dalam kumpulan besar teks. Dalam *text retrieval*, pencarian representasi suatu teks pada kamus dikenal dengan *indexing*. Proses dari *indexing* meliputi *tokenization*, *stopword removal* (eliminasi *stopword*), *stemming*, dan *term weight*. Namun pada penelitian ini hanya digunakan *stopword removal* dan *stemming*.

Eliminasi Stopword

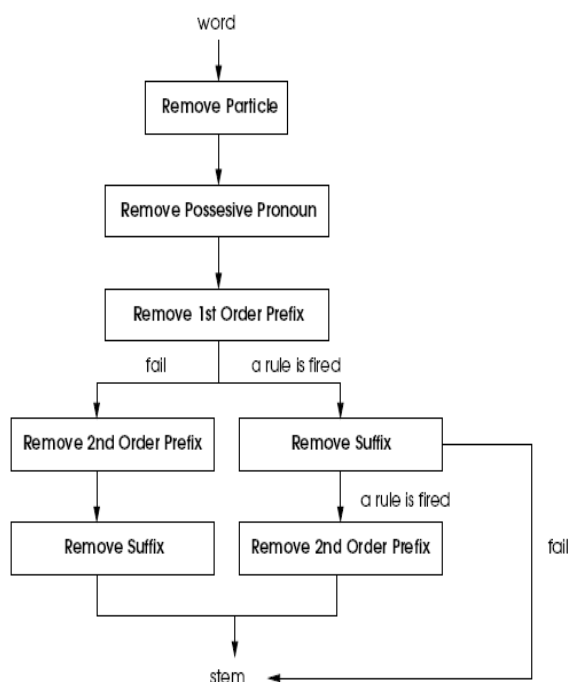
Proses penghilangan tersebut dilakukan dengan cara mencocokkan kata dengan *stoplist*. Jika kata yang dicocokkan ada pada *stoplist*, maka kata tersebut tidak diikuti dalam proses selanjutnya. Contoh kata yang termasuk dalam *stopword* adalah dan, yang, atau, dan lain sebagainya tergantung dari konteks *corpus* yang digunakan sebagai bahan.

Stemming

Stemming adalah proses pemetaan dari penguraian berbagai bentuk kata baik itu prefik, sufik, maupun gabungan antara prefik dan sufik, menjadi bentuk kata dasarnya (*stem*). Algoritma *Stemming* yang digunakan adalah Potter Stemmer untuk Bahasa Indonesia. Ilustrasi algoritmanya dapat dilihat pada Gambar 3.

Algoritma *stemming* pada Gambar 3 tersebut dapat mengatasi akhiran yang berupa partikel, akhiran yang menunjukkan kata ganti kepemilikan, prefiks (imbuan), sufiks (akhirian), dan gabungan antara prefiks dan sufiks. Sebelum melakukan proses penghilangan imbuan pada Potter Stemmer, dilakukan perhitungan *measure*. Hal ini untuk menanggulangi *conflation* yang dilakukan pada kata dasar. Algoritma Porter Stemmer ini dipilih karena mempunyai struktur morfologi yang sama dengan

Bahasa Indonesia yang tersusun dari kombinasi sufiks dan atau prefiks.



Gambar 3. Ilustrasi Potter Stemmer [2]

5. Vector Space Model

Untuk pembuatan kemiripan Alkitab digunakan teori *vector space model*. *Vector space model* ini biasanya digunakan dalam aplikasi search engine seperti Google. Dalam *vector space model* terdapat perhitungan *tf* (*term frequency*) dan *idf* (*inverse document frequency*). Perhitungan *tf* tersebut dapat dilakukan dengan menggunakan salah satu dari rumus di bawah ini.

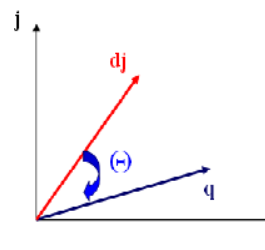
1. $tf = f_{d,t} / \max(f_{d,t})$
2. $tf = f_{d,t} / \sum f_{d,t}$
3. $tf = \begin{cases} 0, & \text{if } (d, t) = 0 \\ 1 + \log(1 + \log(n(d, t))), & \text{otherwise} \end{cases}$

Sedangkan *inverse document frequency* berfungsi untuk menurunkan nilai kordinat dari term-term yang terdapat pada banyak dokumen. Karena tidak semua sumbu pada *vector space* sama pentingnya, terdapat juga dokumen yang memiliki *noise*. Variasi perhitungan *idf* seperti tampak di bawah ini:

1. $idf = 1 / f_t$
2. $idf = \log_2(1 + N / f_t)$

$$3. idf = \log_2((N - f_t)/f_t)$$

Dari perhitungan di atas dapat dihitung nilai *similarity*-nya. Pada *vector space* ini dokumen-dokumen direpresentasikan dengan *vector-vector* pada ruang multidimensi, yang mana diilustrasikan pada Gambar 2 dimensi seperti pada Gambar 4.



Gambar 4. Ilustrasi *Similarity* 2 Dimensi

Dari Gambar 4, *vector* *dj* merupakan dokumen yang digunakan sebagai pencarian. Sedangkan *q* adalah *query* pencariannya. Untuk menentukan tingkat keterkaitannya maka dihitung sudutnya dengan menggunakan *cosinus*.

$$Similarity(D_i, D_j) = \frac{\sum_{k=1}^n t_{ik} * t_{jk}}{\sqrt{\sum_{k=1}^n t_{ik}^2} \times \sqrt{\sum_{k=1}^n t_{jk}^2}} \quad (2)$$

6. Proses Penomoran *Strong*

Kegiatan utama pada pembuatan aplikasi ini adalah melakukan proses penomoran *strong* terhadap semua kata atau frase pada Alkitab Perjanjian Baru Bahasa Indonesia sehingga didapatkan nomor *strong*-nya, yang mana nomor *strong* tersebut didapatkan dari Alkitab Perjanjian Baru Bahasa Inggris. Nomor *strong* yang didapatkan tersebut nantinya digunakan dalam proses pencarian kemiripan berdasarkan perikop yang diinputkan. Ada 5 tahap dalam proses penomoran *strong*. Masing-masing tahap akan dijelaskan dibawah.

1. Dengan pendekatan pencarian frekuensi

Pada tahap pertama ini dilakukan dengan cara menghitung frekuensi kemunculan kata pada ayat yang ditemukan nomor *strong*-nya tersebut. Setelah dihitung kemunculannya maka dilakukan *stopword removal* kemudian dilakukan *filtering* dengan *threshold* tertentu. *Stopword removal* dilakukan karena kata yang terdapat pada *stoplist* sering muncul dalam ayat pada Alkitab Perjanjian Baru Bahasa Indonesia. *Threshold* yang telah diuji coba dan dianggap baik adalah minimal nomor *strong* ditemukan pada dua ayat dan memiliki frekuensi

lebih besar sama dengan 0,6.

2. Dengan *word alignment*

Pada tahap kedua ini dilakukan perhitungan frekuensi dengan bantuan *mutual information* yang diharapkan hasilnya dapat lebih baik dari pada hasil pada tahap pertama. Cara kerjanya hampir sama dengan tahap pertama yaitu pada awalnya disediakan dua *corpus*. Perbedaannya terdapat pada cara perhitungan frekuensinya.

$$\begin{aligned}
 X^2 &= \frac{7955 [(69 \times 7882) - (0 \cdot 4)]^2}{(69+0)(69+4)(0+7882)(4+7882)} \\
 &= \frac{7955 [543858 + 4]^2}{69 \times 73 \times 7882 \times 7886} \\
 &= \frac{7955 \times 295785875044}{313087085724} \\
 &= \mathbf{7515,406234}
 \end{aligned}$$

Gambar 5. Contoh Perhitungan MI

Perhitungan tersebut dilakukan dengan cara membuat matriks yang berukuran 2x2. Contoh matriks berukuran 2x2 tersebut dapat dilihat pada Gambar 5. Dari nilai *mutual information* yang didapat tersebut, semakin besar nilainya maka kandidat tersebut dicurigai sebagai kandidat yang utama.

3. Dengan n-gram

Tahap ketiga adalah dengan mengkombinasikan 2 kata. Kombinasi tersebut didasarkan pada teori n-gram. Dari kombinasi 2 kata tersebut didapatkan kandidat yang nantinya dilakukan *filtering* dengan syarat nilai *mutual information* yang pertama lebih besar dari nilai *mutual information* yang kedua. Dan nilai O_{11} lebih besar dari 1 dan nilai O_{11} lebih besar dari O_{12} .

4. Dengan *stemmed corpus*

Tahap keempat adalah melakukan *stemming* pada *corpus* Alkitab Perjanjian Baru Bahasa Indonesia. Tahap ini diperlukan karena perbedaan antara tata bahasa Inggris dan tata bahasa Indonesia yang dapat membuat frekuensi berbeda untuk kandidat nomor *strong* dan kata-kata. *Corpus* hasil *stemming* tersebut akan digunakan sebagai *corpus* dalam proses tahap 1-3.

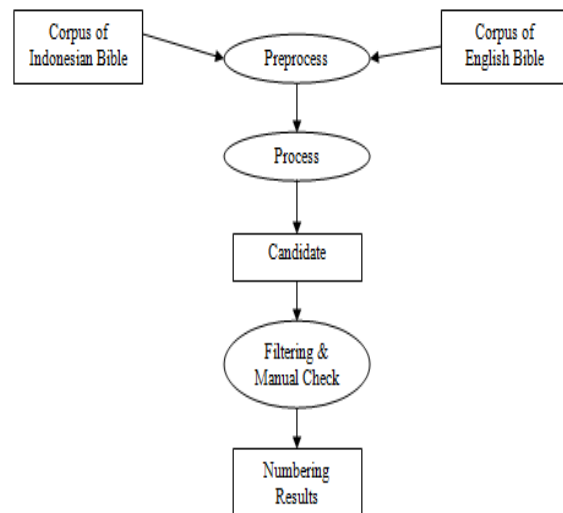
5. Dengan pencarian kembali kandidat *strong* pada tahap 1

Tahap kelima adalah pencarian kembali kandidat *strong* pada tahap I, yang *threshold* kandidat pertama lebih besar dari pada *threshold* kandidat kedua. Selain itu dilakukan pencarian kandidat berdasarkan *proper name*, dan nomor *strong* yang termasuk dalam konjungsi, preposisi, dan kata ganti.

7. Arsitektur Sistem

Sasaran pembuatan penelitian ini adalah membuat

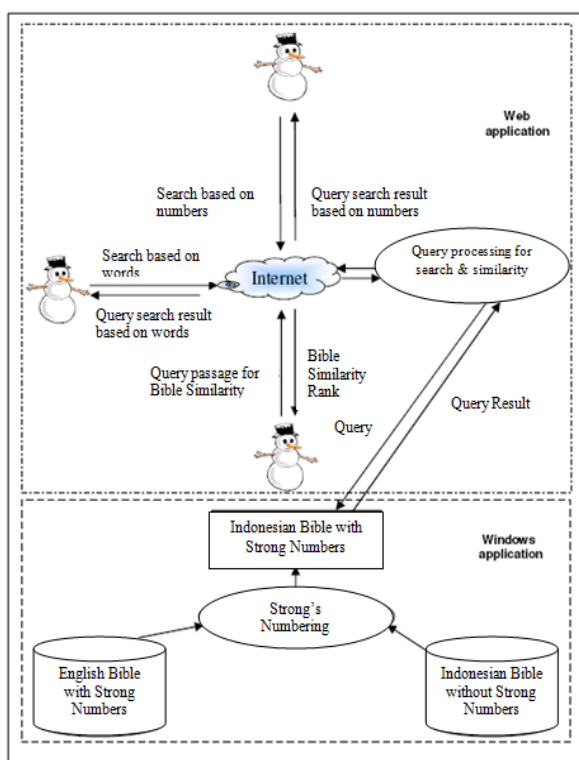
aplikasi penomoran *strong* pada Alkitab Perjanjian Baru Bahasa Indonesia. Hasil dari Alkitab Perjanjian Baru Bahasa Indonesia dengan nomor *strong* ini dapat digunakan untuk pembuatan fitur lainnya yaitu *similarity*. Arsitektur umum untuk penomoran *strong* dari 5 tahap tersebut adalah preproses, proses, kandidat, *filtering* dan cek manual. Arsitektur untuk aplikasi *desktop* dapat dilihat pada Gambar 6.



Gambar 6. Arsitektur untuk Aplikasi Desktop

Gambar 6 menjelaskan proses untuk mendapatkan nomor *strong*. Tahap pertama, Alkitab Perjanjian Baru Bahasa Indonesia dan Inggris diproses dengan menggunakan eliminasi *stopword* atau *stemming*. Ini dapat menghapus semua konjungsi. Tahap kedua, dari hasil dari tahap *preprocessing* lakukan 5 tahap proses penomoran *strong*. Tahap ketiga, kita akan mendapatkan kandidat dari tahap kedua. Pada tahap keempat, kita harus menyaring kandidat dengan menggunakan batas *threshold* yang telah ditentukan sebelumnya, dan secara manual memeriksa nomor *strong* yang memiliki frekuensi sama dengan 1. Pada tahap terakhir, kita akan mendapatkan nomor-nomor *strong*.

Arsitektur untuk pencarian *similarity* dan Alkitab berbasis situs adalah input yang diberikan oleh pengguna akan dikirimkan melalui internet, kemudian dilakukan pengolahan *query* berdasarkan input dari pengguna, baik untuk fitur pencarian dan fitur kesamaan (*similarity*). Data yang digunakan dalam aplikasi ini diperoleh dari nomor *strong* pada aplikasi yang dibuat sebelumnya. Dari hasil *query*, pengolahan akan diserahkan kepada pengguna melalui internet. *Output* untuk permintaan kesamaan ini memuat daftar yang paling mirip di bagian ini diurutkan secara *descending* (menurun). Arsitektur untuk aplikasi situs ditampilkan pada Gambar 7.



Gambar 7. Arsitektur untuk Aplikasi Situs

8. Uji Coba

Aplikasi penomoran *strong* dengan lima tahap yang dilakukan dalam penelitian ini telah dapat memberikan penomoran *strong* yang tepat pada Alkitab Bahasa Indonesia. Hal ini dapat diketahui dari penanganan penomoran *strong* sampai di belakang kata tersebut. Dengan metode penyusunan yang dilakukan didapatkan Alkitab Perjanjian Baru Bahasa Indonesia dengan nomor *strong* dengan tingkat keberhasilan 61,03%.

Stemming dapat digunakan untuk membantu meningkatkan akurasi penomoran. Hal ini dapat dilihat bahwa *stemming* dapat meningkatkan 7,50% hasil penomoran *strong* dari penomoran *strong* tanpa *stemming*.

Kesulitan terutama untuk melakukan penomoran *strong* sampai akurasi yang sempurna adalah fakta bahwa terdapat 1839 nomor *strong* atau 33,29% yang hanya dipakai satu kali dalam Alkitab Perjanjian Baru.

Uji coba dilakukan dengan beberapa jenis kasus dan dapat dilihat pada Tabel 1 berikut ini.

Dari hasil uji coba Tabel 1 dapat dilihat bahwa penomoran *strong* tidak dapat mencapai target yang ditentukan dikarenakan alasan yang telah disebutkan. Namun dengan adanya penomoran *strong* ini,

pengguna dapat melakukan pembelajaran Alkitab dengan lebih mudah dan efisien. Selain itu dapat digunakan untuk mengetahui tingkat keterkaitan antar perikop yang ada pada Alkitab Perjanjian Baru Bahasa Indonesia.

Tabel 1. Uji Coba Penomoran *Strong*

| Jenis Uji Coba | Output Uji Coba | Alasan |
|--|---|--|
| Penomoran <i>strong</i> yang hanya memiliki satu frekuensi | Arti kata yang memiliki satu frekuensi tidak berhasil dinomori | Hanya dapat dilakukan untuk frekuensi lebih dari atau sama dengan 2. Karena dengan threshold minimal 1 ayat tidak dapat dipastikan kebenaran kandidatnya |
| Penomoran <i>strong</i> untuk ejaannya tidak sama | Arti kata yang berbeda ejaannya tidak berhasil dinomori | Penghitungan kandidat didasarkan pada jumlah frekuensi kemunculan kandidat kata tanpa dilakukan pengecekan ejaannya |
| Penomoran <i>strong</i> yang memiliki imbuhan | Arti kata dapat dinomori | Dilakukan <i>stemming</i> sehingga didapatkan root katanya |
| Penomoran <i>strong</i> dimana terdapat perbedaan arti kata | Dapat dinomori tetapi tidak 100% dapat dinomori | Kandidat penomoran <i>strong</i> yang diambil hanya yang kandidat <i>strong</i> dengan nilai frekuensi atau mutual information yang tertinggi |
| Penomoran <i>strong</i> untuk jenis kata proper name | Hampir semua penomoran <i>strong</i> dengan jenis kata proper name dapat ditangani kecuali untuk nomor <i>strong</i> yang hanya memiliki satu frekuensi | Karena kemunculan nomor <i>strong</i> dapat dipastikan terletak pada ayat-ayat tertentu dan memiliki arti yang sama, sehingga kandidat arti kata tersebut lebih kuat |
| Penomoran <i>strong</i> untuk frase | Dapat ditangani namun tidak 100% dapat dinomori | Kandidat frase yang benar tersebut tidak memiliki nilai mutual information yang tertinggi |
| Penomoran <i>strong</i> untuk jumlah kata yang berbeda antara Alkitab Bahasa Indonesia dengan Alkitab Bahasa Inggris | Penomoran <i>strong</i> dengan jumlah kata yang berbeda dapat dinomori sesuai dengan nilai mutual information | Penomoran tersebut tidak tergantung dari banyaknya jumlah kata, tetapi tergantung pada nilai frekuensi atau nilai mutual informationnya |

Modul *similarity* adalah modul yang digunakan untuk mencari kemiripan perikop pada Alkitab Perjanjian Baru. Pencarian kemiripan perikop

tersebut dilakukan dengan cara memilih perikop yang hendak dicari kemiripannya dalam daftar perikop. Hasil dari pencarian kemiripan perikop tersebut dapat dilihat pada Gambar 8. Pada gambar tersebut *query* yang diinputkan oleh pengguna adalah Yohanes pembaptis yang terdapat pada injil Mrk 1:1-8. Hasil dari pencarian kemiripan perikop pada sistem adalah injil Mat 3:1-12. Hasil tersebut sama dengan kemiripan yang terdapat pada Alkitab yang dapat dilihat di bawah judul perikopnya.



Gambar 8. Hasil Uji Coba Modul *Similarity*

9. Kesimpulan

Dengan adanya Alkitab Perjanjian Baru Bahasa Indonesia yang dilengkapi dengan nomor *strong*, pembelajaran Alkitab menjadi lebih mudah dan tersedia secara *online*. Karena dari nomor *strong* tersebut dapat diketahui asal katanya sehingga dapat mengurangi kesalahpahaman dalam pentafsiran. Kesimpulan lain yang dapat diambil dari hasil pengamatan pada setiap tahap penomoran *strong* yang dilakukan pada pembuatan dan uji coba aplikasi sebagai berikut.

1. *Word Alignment* dapat membantu menyelesaikan permasalahan penomoran *strong*, dengan cara mengetahui keterkaitan antara nomor *strong* pada Alkitab Bahasa Inggris dengan kata pada Alkitab Bahasa Indonesia.
2. Penyelesaian masalah penomoran *strong* untuk frase yang menimbulkan kesulitan tersendiri dapat diselesaikan dengan menggunakan perhitungan *n-gram*, karena cara ini dapat mencari keterkaitan antara dua atau lebih kata.
3. Manfaat nyata penomoran *strong* pada aplikasi *web* adalah *similaritas* dan *search* pada Alkitab yang dapat dilakukan dengan melakukan pencarian kata berdasarkan nomor *strong*-nya.

4. Walaupun *stopword* memiliki nomor *strong*, dalam perhitungan hasil *search* dan *similaritas*, tetap akan menyebabkan tidak akurat.

Selain keempat hal tersebut, program ini terbukti dapat menyusun penomoran *strong* pada Alkitab Perjanjian Baru berbahasa Indonesia yang sebelumnya ada pada Alkitab bahasa Inggris ke bahasa Yunani.

REFERENSI

- [1] Bloomfield, Maurice, *A Vedic Concordance*. Motilal Banarsidass Publ. 1990
- [2] Z. Talla, Fadillah., "A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia". <http://info.science.uva.nl/pub/theory/illc/research-reports/MoL-2003-02.text.pdf>. 2003.
- [3] F.Brown, Peter., "A Statistical Approach to Machine Translation". www.aclweb.org/anthology-new/J/J90-2002.pdf. 2002.