

IMPLEMENTASI *LEVENSHTEIN DISTANCE*
UNTUK SISTEM PENGHASIL SOAL *ERROR IDENTIFICATION*
DALAM TOEFL

Oleh
Faisal Syaiful Anwar — faisalsyfl@gmail.com
1403356

ABSTRAK

Test of English as a Foreign Language (TOEFL) merupakan salah satu evaluasi yang membutuhkan kualitas soal yang mumpuni. Saat ini, tidak dapat dimungkiri bahwa pembuatan soal evaluasi benar-benar memakan waktu bagi para pembuat soal. Penggunaan teknologi komputer tentunya sangat mengurangi waktu yang dihabiskan pembuat soal dalam membuat soal evaluasi. Sehingga, soal yang dihasilkan bisa secara cepat dan banyak. Untuk itu, penelitian ini membuat sebuah model dan implementasi sistem untuk menghasilkan tipe soal *error identification* secara otomatis dari berita. Soal dihasilkan dari kalimat-kalimat pada situs berita asing yang telah dipraproses dan diukur kesamaan *part of speech* dengan 3024 soal dari *ebook* sumber, menggunakan Levenshtein Distance. Model sistem ini memiliki beberapa tahapan yaitu: (1) Praproses (2) *Part of speech tagging*. (3) POS *similarity*. (4) *Ranking*. (5) Penentuan *underline*. (6) Pemilihan pengecoh. Penelitian ini menghasilkan sebuah sistem yang dapat menghasilkan soal *error identification* secara otomatis. Selain itu, hasil analisis soal oleh sistem menunjukkan persentase 45% kebenaran soal. Sedangkan, hasil analisis evaluasi kualitas soal oleh 2 orang *expert* menunjukkan persentase 82%. Sehingga, kualitas soal yang dihasilkan termasuk dalam kategori sangat baik.

Kata kunci: TOEFL, Penghasil Soal Otomatis, *Natural Language Processing*, *Levenshtein Distance*

**IMPLEMENTASI LEVENSHTTEIN DISTANCE
UNTUK SISTEM PENGHASIL SOAL ERROR IDENTIFICATION
DALAM TOEFL**

Arranged by

Faisal Syaiful Anwar — faisalsyfl@gmail.com

1403356

ABSTRACT

Test of English as a Foreign Language (TOEFL) is one of the evaluations that requires good quality. At this time, it can not be denied that making the evaluation question is really time consuming for the problem makers. The use of computer technology certainly greatly reduces the time spent by the problem maker in making the evaluation question. For that reason, this research makes a computation model to generate the type of error identification question automatically. Questions raised from sentences on foreign news sites that have been processed and measured the similarity of part of speech with 3024 matter of training data from source ebook, using Levenshtein Distance. This system model has several stages: (1) Preprocessing. (2) Part of speech tagging. (3) POS similarity. (4) Ranking. (5) Determination of the underline. (6) Distractor Selection. This research produces a system that can generate the question of error identification automatically. In addition, the result of quality evaluation analysis questioned by 2 experts showed 82% percentage. Thus, the quality of the questions is very good.

Keyword: TOEFL, automatic question generation, Natural Language Processing, Levenshtein Distance.

Faisal Syaiful Anwar, 2017

IMPLEMENTASI LEVENSHTTEIN DISTANCE

UNTUK SISTEM PENGHASIL SOAL ERROR IDENTIFICATION DALAM TOEFL

universitas Pendidikan Indonesiarepository.upi.eduperpustakaan.upi.edu