

## Perbaikan Sistem *Stanford RTE* pada Kalimat Mengandung Ekspresi Aritmatika

Rakhmat Arianto<sup>1</sup>, Daniel Oranova Siahaan<sup>2</sup>, Ahmad Saikhu<sup>3</sup>

<sup>1,2,3</sup>Institut Teknologi Sepuluh Nopember, Surabaya

e-mail: anto.it05@gmail.com<sup>1</sup>, daniel@if.its.ac.id<sup>2</sup>, saikhu@if.its.ac.id<sup>3</sup>

### Abstrak

*Sistem Stanford Recognizing Textual Entailment adalah sistem yang dapat mendeteksi entailment maupun kontradiksi yang terkandung dalam pasangan kalimat Text dan Hypothesis. Pada tahun 2009, dilakukan penelitian pengembangan sistem Stanford RTE dengan menggabungkan sistem Stanford RTE pada tahun 2006 dengan sistem Stanford RTE pada tahun 2008. Salah satu kelemahan pada sistem Stanford RTE tahun 2009 adalah kesalahan deteksi kontradiksi pada pasangan kalimat Text dan Hypothesis mengandung ekspresi aritmatika. Untuk menutupi kelemahan tersebut, ditambahkan fitur aritmatika dalam sistem Stanford RTE yang khusus memproses pasangan kalimat Text dan Hypothesis mengandung ekspresi aritmatika. Fitur aritmatika dibangun dengan empat tahap utama, yaitu tahap analisa linguistik, tahap pencarian nilai kemiripan kata, tahap penentuan operator aritmatika, dan tahap penyimpulan entailment, kontradiksi, atau tidak diketahui. Fitur aritmatika telah dilakukan pengujian terhadap 30 pasang kalimat Text dan Hypothesis mengandung ekspresi aritmatika yang diambil dari halaman website berita dengan hasil tingkat keberhasilan mencapai 80%.*

**Kata kunci:** ekspresi aritmatika, entailment, kontradiksi, Stanford RTE

### Abstract

*Stanford Recognizing Textual entailment system is a system that detects entailment or contradiction is contained in a sentence pair Text and Hypothesis. In 2009, research conducted Stanford RTE system development by combining Stanford RTE system in 2006 with Stanford RTE system in 2008. One disadvantage of the Stanford RTE system in 2009 was a mistake to pair sentences contradiction detection Text and Hypothesis containing arithmetic expressions. To cover these weaknesses, added features Stanford RTE system of arithmetic in which specialized processing Text and Hypothesis couple sentences containing arithmetic expressions. Features arithmetic built with four main stages, namely the stage of linguistic analysis, word similarity value of finding stage, the stage of determining the arithmetic operators and inference stage entailment, contradiction, or unknown. Features arithmetic has been tested on 30 pairs of sentences containing the Text and Hypothesis arithmetic expressions taken from news web pages with the results achieved 80% success rate.*

**Keywords:** arithmetic expression, contradiction, entailment, Stanford RTE

### 1. Pendahuluan

*Stanford Recognizing Textual Entailment (Stanford RTE)* merupakan sistem yang dibangun bertujuan untuk mendeteksi pasangan kalimat mengandung kontradiksi, *entailment*, atau tidak diketahui. Pasangan kalimat yang digunakan masukan terdiri dari kalimat *Text* yang berisikan kalimat uraian dan kalimat *Hypothesis* yang merupakan kesimpulan dari kalimat *Text*. Arti kata kontradiksi menurut Kamus Besar Bahasa Indonesia adalah pertentangan antara dua hal yang sangat berlawanan atau bertentangan. Sedangkan arti kata *entailment* menurut WordNet 3.0 adalah sesuatu yang disimpulkan (dideduksi atau terkandung atau tersirat).

*Stanford RTE* pertama kali dibangun pada tahun 2006 [1], bertujuan untuk mendeteksi pasangan kalimat *Text* dan *Hypothesis* yang mengandung *entailment*. Penelitian selanjutnya yang dilakukan untuk mengembangkan sistem *Stanford RTE* pada tahun 2008 [2] dimana sistem

*Stanford* RTE yang awalnya digunakan untuk mendeteksi pasangan kalimat *Text* dan *Hypothesis* yang mengandung *entailment*, digunakan untuk mendeteksi pasangan kalimat *Text* dan *Hypothesis* yang mengandung kontradiksi. Pada tahun 2009, dilakukan pengembangan terhadap sistem *Stanford* RTE berdasarkan hasil penelitian tahun 2008 yang memiliki tingkat akurasi rendah, maka dilakukan pengembangan dengan cara menggabungkan sistem *Stanford* RTE pada tahun 2006 dengan sistem *Stanford* RTE pada tahun 2008 [3].

Pada penelitian yang dilakukan tahun 2009, sistem *Stanford* RTE yang telah dilakukan perbaikan masih memiliki beberapa kelemahan. Salah satu kelemahan tersebut adalah kesalahan deteksi kontradiksi terhadap pasangan kalimat *Text* dan *Hypothesis* yang mengandung ekspresi aritmatika. Contoh pasangan kalimat *Text* dan *Hypothesis* mengandung ekspresi aritmatika ditunjukkan pada Tabel 1.

**Tabel 1. Pasangan kalimat mengandung ekspresi aritmatika**

Text	Hypothesis
Qasab and an accomplice carried out the assault on the main railway station, killing all <u>52 people</u> in the luxury hotel and <u>9 employers</u> in Jewish cultural center.	The assault on the main railway station, killing <u>61 people</u> .

Pasangan kalimat *Text* dan *Hypothesis* pada Tabel 1 merupakan pasangan kalimat mengandung ekspresi aritmatika karena berdasarkan pemaknaan manusia, frase *52 people* dan *9 employers* apabila dibandingkan dengan frase *61 people* memiliki makna pernyataan yang benar sehingga pasangan kalimat *Text* dan *Hypothesis* pada Tabel 1 juga termasuk dalam pasangan kalimat *entailment*. Namun apabila menggunakan sistem *Stanford* RTE, pasangan kalimat *Text* dan *Hypothesis* pada Tabel 1 akan terdeteksi sebagai pasangan kalimat kontradiksi karena sistem *Stanford* RTE pada proses ekstraksi fitur kontradiksi akan memasukkan pasangan kalimat tersebut pada fitur perbedaan angka sehingga sistem *Stanford* RTE akan secara langsung membandingkan frase *52 people* dengan *61 people* dan *9 employers* dengan *61 people*. Hasil perbandingan angka yang dilakukan sistem *Stanford* RTE tidak mempunyai makna pernyataan yang benar sehingga pasangan kalimat tersebut termasuk dalam pasangan kalimat kontradiksi.

Untuk menutupi kesalahan deteksi kontradiksi sistem *Stanford* RTE terhadap pasangan kalimat *Text* dan *Hypothesis* mengandung ekspresi aritmatika, ditambahkan sebuah fitur aritmatika pada sistem *Stanford* RTE. Fitur aritmatika dibangun dengan empat tahap utama, yaitu tahap analisa linguistik menggunakan *Stanford CoreNLP*[4], tahap pencarian nilai kemiripan kata menggunakan *WordNet Similarity*[5], tahap penentuan operator aritmatika berdasarkan makna kata [6] dan tahap penyimpulan sebagai pasangan kalimat *entailment*, kontradiksi, atau tidak diketahui. Hasil pengujian fitur aritmatika terhadap 30 pasang kalimat *Text* dan *Hypothesis* mengandung ekspresi aritmatika yang diambil dari kalimat berita mendapatkan tingkat keberhasilan mencapai 80%.

## 2. Metodologi Penelitian

Tahapan dalam penelitian ini, diawali dengan pengkajian terhadap pustaka-pustaka yang mendukung penelitian, penggalan data sebagai studi kasus penelitian, pengembangan metode, pengujian metode yang telah dikembangkan, pembahasan terhadap hasil pengujian, dan memberikan kesimpulan terhadap penelitian yang telah dilakukan.

### 2.1 Penggalan Data

Dalam penelitian sebelumnya, digunakan studi kasus RTE4 [7] yang telah disediakan oleh pihak penyelenggara konferensi. RTE4 merupakan studi kasus yang berisi 1000 pasang kalimat *Text* dan *Hypothesis* mewakili bentuk kalimat tanya-jawab, kalimat informasi, ekstraksi informasi, dan ringkasan dari dokumen. Namun, 1000 pasang kalimat pada RTE4 yang mengandung ekspresi aritmatika hanya terdapat pada ID 332 sehingga diperlukan tambahan

studi kasus yang diambil dari halaman *website* berita sebanyak 30 pasang kalimat. Contoh studi kasus yang digunakan dalam penelitian ini terdapat pada Tabel 2.

**Tabel 2. Studi Kasus Penelitian**

ID	SUMBER	TEXT	HYPOTHESIS
1	<a href="http://edition.cnn.com/2012/10/21/us/george-mcGovern-dead/index.html?hpt=hp_t3">http://edition.cnn.com/2012/10/21/us/george-mcGovern-dead/index.html?hpt=hp_t3</a>	The son of a Methodist minister who was a Republican, McGovern was born in Avon, South Dakota, on July 19, 1922. Six years later, his family moved an hour north to Mitchell, where McGovern graduated from Mitchell High School in 1940.	In 1928, McGovern's family moved an hour north to Mitchell.
.	.	.	.
.	.	.	.
.	.	.	.
30	<a href="http://thedailynews.com/news/article_448d4a62-6cf9-11e2-af8a-0019bb2963f4.html">http://thedailynews.com/news/article_448d4a62-6cf9-11e2-af8a-0019bb2963f4.html</a>	Genesee County's December 2012 unemployment rate was 8.2 percent, 0.8 of a percentage point more than it was in November.	Genesee County's unemployment rate was 7.3 percent in November.

## 2.2 Pengembangan Metode

Pengembangan metode yang dilakukan adalah sistem *Stanford RTE* tahun 2009 [3] ditambahkan fitur aritmatika yang berfungsi untuk memproses pasangan kalimat *Text* dan *Hypothesis* mengandung ekspresi aritmatika. Pengembangan metode dilakukan pada proses penyimpulan kontradiksi dimana didalamnya terdapat proses ekstraksi fitur kontradiksi. Fitur aritmatika diletakkan setelah fitur perbedaan numerik, tanggal, atau waktu pada proses ekstraksi fitur kontradiksi. Fitur aritmatika dapat memproses pasangan kalimat *Text* dan *Hypothesis* mengandung ekspresi aritmatika pada jenis angka numerik, tanggal, dan persen.

### 2.2.1 Jenis Angka Numerik

Penggunaan fitur aritmatika pada studi kasus pasangan kalimat *Text* dan *Hypothesis* mengandung ekspresi aritmatika pada jenis angka numerik telah dilakukan penelitian tahun 2013 [8]. Apabila kata kerja yang memiliki keterkaitan dengan kata benda sebagai satuan pada kalimat *Text* dan *Hypothesis* sama, maka dilakukan pencarian nilai kemiripan kata dari kata benda yang terkait dengan kata kerja yang sama. Apabila nilai kemiripan kata melebihi batas minimal yang telah ditentukan, maka dilakukan penjumlahan pada angka yang terkait dengan kata benda pada kalimat *Text*. Hasil aritmatika dari kalimat *Text* akan dibandingkan dengan angka dan kata benda pada kalimat *Hypothesis*. Apabila hasil perbandingan angka dan kata benda pada kalimat *Text* dan *Hypothesis* mengandung pernyataan yang benar, maka pasangan kalimat tersebut termasuk dalam pasangan kalimat *entailment*. Apabila hasil perbandingan angka dan kata benda pada kalimat *Text* dan *Hypothesis* mengandung pernyataan yang salah, maka pasangan kalimat tersebut termasuk dalam pasangan kalimat kontradiksi. Apabila pasangan kalimat tidak termasuk dalam pasangan kalimat *entailment* dan kontradiksi, maka pasangan kalimat tersebut termasuk pasangan kalimat tidak diketahui.

### 2.2.2 Jenis Angka Tanggal

Penggunaan fitur aritmatika pada pasangan kalimat *Text* dan *Hypothesis* mengandung ekspresi aritmatika pada jenis angka tanggal, diperlukan analisa terhadap hasil deteksi pada *Normalized NER* dari kalimat *Text* untuk menentukan operator aritmatika yang dilakukan pada kalimat *Text*, pencarian tanggal yang akan dilakukan operasi aritmatika, dan bagian dari tanggal yang akan dilakukan operasi aritmatika apakah pada tanggal, bulan atau tahun.

**Tabel 3. Studi kasus jenis angka tanggal**

ID	TEXT	HYPOTHESIS	HUMAN	STANFORD RTE
1	The son of a Methodist minister who was a Republican, McGovern was born in Avon, South Dakota, on <u>July 19, 1922</u> . <u>Six years later</u> , his family moved an hour north to Mitchell, where McGovern graduated from Mitchell High School in 1940.	In 1928, McGovern's family moved an hour north to Mitchell.	ENTAILMENT	CONTRADICTION

Tabel 3 menunjukkan pasangan kalimat *Text* dan *Hypothesis* mengandung ekspresi aritmatika pada jenis angka tanggal yang ditunjukkan pada frase *July 19, 1922* dan *six years later* dibandingkan dengan angka *1928*. Kolom HUMAN menunjukkan hasil anotasi berdasarkan pemikiran manusia dan kolom *Stanford RTE* merupakan hasil deteksi menggunakan sistem *Stanford RTE* tahun 2009 [3].

**Tabel 4. Deteksi tanggal kalimat pertama pada kalimat *Text***

Kata	NER	Normalized NER
July	DATE	1922-07-19
19	DATE	1922-07-19
,	DATE	1922-07-19
1922	DATE	1922-07-19

**Tabel 5. Deteksi tanggal kalimat kedua pada kalimat *Text***

Kata	NER	Normalized NER
Six	DATE	OFFSET P6Y
years	DATE	OFFSET P6Y
later	DATE	OFFSET P6Y

Tabel 4 menunjukkan hasil deteksi frase pada kalimat pertama dari kalimat *Text* yaitu 1922-07-19. Sedangkan kalimat kedua dari kalimat *Text* pada Tabel 5, deteksi tanggal menunjukkan "OFFSET P6Y". Kata OFFSET mempunyai arti "lebih" sehingga operator aritmatika yang akan dilakukan adalah penjumlahan dan kata "P6Y" (Present 6 Years) mempunyai arti terjadi lebih 6 tahun maka angka 1922-07-19 akan dilakukan penjumlahan 6 tahun sehingga menghasilkan angka 1928-07-19. Hasil penjumlahan pada kalimat *Text* akan dibandingkan dengan hasil deteksi tanggal pada kalimat *Hypothesis* yaitu 1928. Sehingga, pasangan kalimat *Text* dan *Hypothesis* pada Tabel 3 dapat diketahui mengandung pernyataan yang benar dan termasuk pada pasangan kalimat entailment.

### 3. Kajian Pustaka

#### 3.1 *Stanford RTE* Tahun 2006

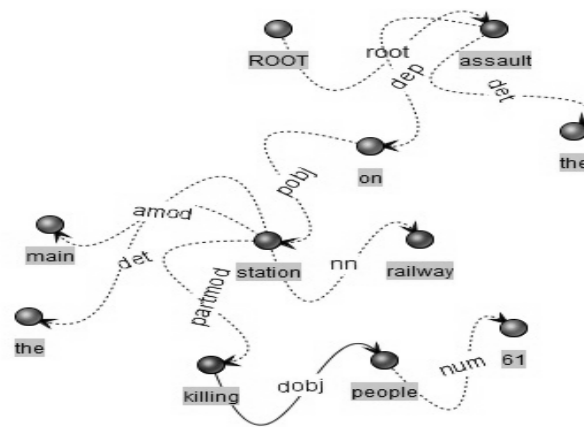
Pada tahun 2006, dilakukan penelitian untuk mengikuti konferensi umum bertajuk *PASCAL Recognizing Textual Entailment*. Dalam penelitian tersebut menghasilkan sebuah sistem *Stanford RTE* yang digunakan untuk mendeteksi pasangan kalimat *Text* dan *Hypothesis* termasuk dalam pasangan kalimat *entailment* atau tidak diketahui. Sistem *Stanford RTE* terdiri dari tiga tahap utama, yaitu tahap analisa linguistik, tahap penyelarasan pohon dependensi antar kata dalam kalimat, dan tahap penyimpulan *entailment* atau tidak diketahui [1].

Tahap analisa linguistik bertujuan untuk mendapatkan informasi semantik sebanyak mungkin dari masing-masing kalimat *Text* dan *Hypothesis*. Analisa yang dilakukan meliputi analisa keterkaitan antar kata dalam kalimat, analisa penamaan entitas dalam kalimat, dan analisa penyederhanaan kolokasi kata yang berdekatan dalam kalimat. Keseluruhan hasil analisa linguistik tersirat dalam pohon dependensi yang dihasilkan *StanfordCoreNLP*. Contoh

dependensi antar kata pada kalimat *the assault on the main railway station, killing 61 people*, ditunjukkan pada Tabel 6 dan bentuk visualisasi dari pohon dependensi ditunjukkan pada Gambar 1 menggunakan alat bantu *Grammarscope*[9].

**Tabel 6. Dependensi antar kata dalam kalimat**

Dependensi Kata
det ( assault-2 , The-1 )
nsubj ( killing-9 , assault-2 )
det ( station-7 , the-4 )
amod ( station-7 , main-5 )
nn ( station-7 , railway-6 )
prep_on ( assault-2 , station-7 )
num ( people-11 , 61-10 )
dobj ( killing-9 , people-11 )



**Gambar 1. Visualisasi dependensi antar kata dalam kalimat**

Tahap penyelarasan dependensi antar kata dalam kalimat bertujuan untuk mendapatkan keselarasan tiap kata dalam kalimat *Text* dan *Hypothesis* berdasarkan pohon dependensi antar kata yang dihasilkan oleh tahap analisa linguistik. Sistem *Stanford RTE* pada tahun 2006 membangun keselarasan kata dengan pemetaan nilai tertinggi dari setiap *node* pohon dependensi antar kata pada kalimat *Hypothesis* dengan *node* tunggal atau dengan *node* kosong dari pohon dependensi antar kata dalam kalimat *Text*. Dalam proses penyelarasan antara pohon dependensi digunakan metode *Stochastic*.

Tahap penyimpulan *entailment* bertujuan untuk menentukan apakah pasangan kalimat *Text* dan *Hypothesis* termasuk dalam pasangan kalimat *entailment* atau tidak diketahui. Tahap penyimpulan dilakukan dengan cara menempatkan hasil tahapan sebelumnya pada model fitur yang telah disediakan. Jenis model fitur meliputi sintaktik, leksikal, dan fenomena semantik termasuk pasangan kalimat *Text* dan *Hypothesis* yang mengandung kata kerja faktif, polaritas, antonim, kata kerja bantu, kuantitas, kecocokan tanggal, waktu dan angka dalam kalimat, komparabilitas dari struktur sintaktik, dan kualitas dari penyelarasan. Apabila salah satu fitur atau lebih menyatakan pernyataan yang benar, maka pasangan kalimat tersebut termasuk dalam pasangan kalimat *entailment*. Jika tidak terdapat pernyataan yang benar, maka pasangan kalimat tersebut termasuk dalam pasangan kalimat yang tidak diketahui.

### 3.2 Stanford RTE Tahun 2008

Pada tahun 2008, dilakukan penelitian terhadap sistem *Stanford RTE* yang dibangun bertujuan untuk mendeteksi *entailment* pada pasangan kalimat *Text* dan *Hypothesis* digunakan untuk mendeteksi kontradiksi pada pasangan kalimat *Text* dan *Hypothesis*[2]. Penelitian deteksi

kontradiksi pada pasangan kalimat *Text* dan *Hypothesis* menggunakan sistem *Stanford RTE* diperlukan perubahan pada proses pemodelan dan proses penyimpulan.

Kemunculan kontradiksi pada pasangan kalimat *Text* dan *Hypothesis* dapat digolongkan menjadi dua kategori, yaitu kalimat kontradiksi yang timbul karena adanya kata antonim, negasi, dan ketidakcocokan angka termasuk juga pada jenis angka tanggal dan waktu serta kemunculan kontradiksi karena penggunaan kata kerja aktif, kata kerja bantu, perbedaan leksikal, perbedaan struktur kalimat, dan perbedaan penggunaan kata fakta yang terjadi di dunia. Contoh kemunculan kontradiksi dalam kalimat ditunjukkan pada Tabel 7 yang diambil dari pustaka.

**Tabel 7. Kategori kemunculan kontradiksi**

No	Type	Text	Hypothesis
1	Antonim	Capital punishment is a catalyst for more crime.	Capital punishment is a deterrent to crime.
2	Negasi	A closely divided Supreme Court said that juries and not judges must impose a death sentence.	The Supreme Court decided that only judges can impose the death sentence
3	Numerik	The tragedy of the explosion in Qana that killed more than 50 civilians has presented Israel with a dilemma.	An investigation into the strike in Qana found 28 confirmed dead thus far.
4	Modal Verbs	Prime Minister John Howard says he will not be swayed by a warning that Australia faces more terrorism attacks unless it withdraws its troops from Iraq.	Australia withdraws from Iraq.
5	Factive Verbs	The bombers had not managed to enter the embassy.	The bombers entered the embassy.
6	Struktur	Jacques Santer succeeded Jacques Delors as president of the European Commission in 1995.	Delors succeeded Santer in the presidency of the European Commission.
7	Leksikal	In the election, Bush called for U.S. troops to be withdrawn from the peacekeeping mission in the Balkans.	He cites such missions as an example of how America must "stay the course."
8	Fakta	Microsoft Israel, one of the first Microsoft branches outside the USA, was founded in 1989.	Microsoft was established in 1989.

Untuk mendeteksi kontradiksi pada pasangan kalimat *Text* dan *Hypothesis* maka sistem *Stanford RTE* pada tahun 2006 dimodifikasi sehingga memiliki empat tahap utama, yaitu tahap analisa linguistik, tahap penyelarasan pohon dependensi kata dalam kalimat, tahap penyaringan pasangan kalimat merujuk pada peristiwa yang sama bertujuan, dan tahap penyimpulan kontradiksi atau tidak diketahui. *Stanford RTE* pada tahun 2008 menunjukkan adanya perbedaan dengan sistem *Stanford RTE* pada tahun 2006 adalah adanya tahap penyaringan pasangan kalimat yang merujuk pada peristiwa yang sama dan tahap penyimpulan kontradiksi. Tahap penyaringan pasangan kalimat bertujuan untuk memastikan bahwa kontradiksi yang muncul dalam kalimat benar-benar terjadi pada pasangan kalimat yang merujuk pada peristiwa yang sama karena apabila perbedaan dalam pasangan kalimat yang membahas peristiwa yang berbeda, tidak termasuk dalam kategori pasangan kalimat kontradiksi. Sedangkan untuk tahap penyimpulan kontradiksi dilakukan dengan mengekstraksi pada fitur polaritas, fitur struktural, fitur antonim, fitur angka, fitur faktif, fitur kata kerja bantu, dan fitur relasional.

### 3.3 Stanford RTE Tahun 2009

Penelitian pada tahun 2009, dilakukan terhadap sistem Stanford RTE pada tahun 2008 yang masih memiliki kekurangan dalam hal ketepatan dalam mendeteksi kontradiksi pada pasangan kalimat *Text* dan *Hypothesis*. Untuk menutupi kelemahan tersebut, dilakukan pengembangan terhadap sistem Stanford RTE dengan menggabungkan deteksi *entailment* pada sistem Stanford RTE 2006 dengan deteksi kontradiksi pada sistem Stanford RTE 2008 sehingga sistem Stanford RTE pada tahun 2009, memiliki lima tahap utama yaitu tahap analisa linguistik, tahap penyelarasan pohon dependensi kata dalam kalimat, tahap penyaringan pasangan kalimat yang merujuk pada peristiwa yang sama, penyimpulan *entailment*, dan penyimpulan kontradiksi [3].

Penggabungan sistem Stanford RTE dimaksudkan agar disaat mendeteksi kontradiksi pada pasangan kalimat *Text* dan *Hypothesis*, merupakan pasangan kalimat yang benar-benar tidak termasuk dalam pasangan kalimat *entailment* sehingga dapat meningkatkan tingkat akurasi deteksi kontradiksi. Apabila pasangan kalimat *Text* dan *Hypothesis* tidak termasuk pasangan kalimat *entailment*, maka akan dideteksi adanya kontradiksi dalam pasangan kalimat, apabila tidak termasuk dalam pasangan kalimat kontradiksi, maka pasangan kalimat tersebut termasuk dalam pasangan kalimat tidak diketahui.

Sistem Stanford RTE pada tahun 2009 memiliki tiga kelemahan, yaitu kesalahan deteksi pada pasangan kalimat *Text* dan *Hypothesis* yang mengandung ambiguitas, memiliki perbedaan kalimat aktif-pasif, dan mengandung ekspresi aritmatika.

### 3.4 Stanford CoreNLP

Dalam penelitian ini, proses analisa linguistik menggunakan alat bantu Stanford CoreNLP[4]. Alat bantu ini merupakan kumpulan metode-metode yang sangat diperlukan dalam pemrosesan bahasa alami. Alat bantu ini dapat memberikan bentuk dasar dari tiap kata dalam kalimat, jenis tiap kata dalam kalimat, penamaan entitas dalam kalimat, menormalisasi kata yang menunjukkan tanggal, waktu, dan numerik, menandai atas struktur kalimat dalam bentuk frasa dan dependensi kata, serta menunjukkan frase kata benda yang mengacu pada entitas yang sama.

#### 3.4.1 POS Tagger

Stanford CoreNLP dapat memberikan informasi jenis kata dari tiap kata dalam kalimat dikarenakan Stanford CoreNLP memiliki metode POS Tagger didalamnya. Metode ini akan memberikan informasi jenis kata benda, kata kerja, atau yang lainnya pada setiap kata dalam kalimat [10]. Contoh deteksi POS Tagger yang dilakukan pada kalimat *the assault on the main railway station, killing 61 people* ditunjukkan pada Tabel 8.

**Tabel 8. Hasil Penggunaan POS Tagger**

Kata	POS
The	DT
assault	NN
on	IN
the	DT
main	JJ
railway	NN
station	NN
,	,
killing	VBG
61	CD
people	NNS

Hasil deteksi POS Tagger yang ditunjukkan pada Tabel 8, tiap kata memiliki jenis masing-masing, dimana DT berarti *determiner*, NN berarti *Noun*, IN berarti *preposition*, JJ berarti *Adjective*, VBG berarti *verb, gerund or present participle*, CD berarti *Cardinal Number*, NNS berarti *noun, plural*. Keterangan lebih mendetail tentang POS Tagger dapat dilihat pada pustaka.

### 3.4.2 Penamaan Entitas

Sistem *Stanford CoreNLP* dapat menentukan penamaan entitas yang terdapat dalam kalimat dikarenakan di dalam sistem *Stanford CoreNLP* digunakan metode *Stanford NER (Named Entity Recognizer)*. Metode ini juga dikenal sebagai *CRF Classifier* yang berfungsi untuk memberikan penamaan terhadap rangkaian kata dalam kalimat yang menunjukkan nama orang, perusahaan, numerik, tanggal, dan lain-lain [11]. Penamaan entitas banyak digunakan dalam aplikasi berbasis pemrosesan bahasa alami seperti teks dalam bentuk pertanyaan dan jawaban, rangkuman dan sistem dialog. Contoh penamaan entitas dalam kalimat *assault on the main railway station, killing 61 people*, ditunjukkan pada Tabel 9.

**Tabel 9. Hasil penamaan entitas dalam kalimat**

Kata	NER	Normalized NER
The	O	
assault	O	
on	O	
the	O	
main	O	
railway	O	
station	O	
,	,	
killing	O	
61	NUMBER	61.0
people	O	

Hasil penamaan entitas dalam kalimat yang ditunjukkan pada Tabel 9, kata *61* terdeteksi sebagai entitas jenis NUMBER dan hasil setelah dilakukan normalisasi menjadi 61.0 sedangkan untuk kata yang lain tidak menunjukkan entitas sehingga mendapatkan hasil "O". Perbedaan antara NER dengan *Normalized NER* akan terlihat apabila dalam kalimat terdapat frase yang menunjukkan entitas jenis angka tanggal.

### 3.5 WordNet Similarity

Metode ini digunakan untuk mendapatkan nilai kemiripan dan keterkaitan antar kata dengan menggunakan isi dan struktur dari *WordNet*[5]. Pengukuran kemiripan kata menggunakan informasi hierarki dari kata atau *synsets* dan menghitung seberapa tingkat kemiripan kata dengan kata. Misalnya, dalam pengukuran tingkat kemiripan kata *automobile* akan menunjukkan tingkat kemiripan lebih tinggi pada kata *boat* daripada tingkat kemiripan dengan kata *tree* dikarenakan fakta yang menunjukkan bahwa kata *automobile* dan *boat* termasuk kata yang memiliki hierarki di bawah kata *vehicle* dalam hierarki kata benda *WordNet*.

Pengukuran tingkat kemiripan memiliki dua dasar metode, yaitu berdasarkan isi informasi dari jarak terdekat antar kata (*least common subsumer/ LCS*) dan berdasarkan panjang jalur antar kata yang dibandingkan. Salah satu metode yang berdasarkan panjang jalur antar kata adalah *Wu & Palmer (wup)*. Metode *Wu & Palmer* bekerja dengan cara mencari kedalaman LCS dari kata dan menghitung penjumlahan kedalaman dari masing-masing kata. Jarak kedalaman kata



adalah jarak dari kata menuju ke kata yang berada pada simpul akar. Jalur pengukuran antar kata adalah kebalikan arah dari jarak terdekat antar dua kata.

### 3.6 Kata Bermakna Operasi Aritmatika

Dalam pasangan kalimat *Text* dan *Hypothesis* mengandung ekspresi aritmatika, selalu terdapat kata kunci yang mempunyai makna sebagai operator aritmatika. Operator aritmatika nantinya akan menentukan apakah operasi aritmatika yang akan dilakukan baik penjumlahan, perkalian, pembagian, dan pengurangan. Berikut contoh kata kunci yang menunjukkan operator aritmatika pada Tabel 10 [6].

**Tabel 10. Kata kunci bermakna operator aritmatika**

Kata Kunci	Operator Aritmatika
increase	Penjumlahan
sum	Penjumlahan
total	Penjumlahan
added	Penjumlahan
all	Penjumlahan
difference	Pengurangan
left	Pengurangan
fewer	Pengurangan
minus	Pengurangan
reduce	Pengurangan
times	Perkalian
product	Perkalian
at	Perkalian
per	Perkalian
total of	Perkalian
twice	Perkalian
quotient	Pembagian
ratio	Pembagian
per	Pembagian
percent	Pembagian
half	Pembagian

## 4. Hasil dan Pembahasan

Hasil pengujian fitur aritmatika terhadap studi kasus 30 pasang kalimat *Text* dan *Hypothesis* pada jenis angka numerik, tanggal, dan persen ditunjukkan pada Tabel 11.

**Tabel 11. Hasil pengujian fitur aritmatika**

ID	HUMAN	STANFORD RTE	FITUR ARITMATIKA
1.	ENTAILMENT	CONTRADICTION	ENTAILMENT
2.	CONTRADICTION	CONTRADICTION	CONTRADICTION
3.	ENTAILMENT	CONTRADICTION	ENTAILMENT
4.	CONTRADICTION	CONTRADICTION	CONTRADICTION
5.	ENTAILMENT	CONTRADICTION	ENTAILMENT
6.	CONTRADICTION	CONTRADICTION	CONTRADICTION
7.	ENTAILMENT	ENTAILMENT	ENTAILMENT
8.	CONTRADICTION	ENTAILMENT	CONTRADICTION
9.	ENTAILMENT	CONTRADICTION	ENTAILMENT
10.	CONTRADICTION	CONTRADICTION	CONTRADICTION
11.	ENTAILMENT	CONTRACIDTION	ENTAILMENT
12.	CONTRADICTION	CONTRADICTION	CONTRADICTION

13.	ENTAILMENT	ENTAILMENT	UNKNOWN
14.	CONTRADICTION	CONTRADICTION	UNKNOWN
15.	ENTAILMENT	CONTRADICTION	ENTAILMENT
16.	CONTRADICTION	CONTRADICTION	CONTRADICTION
17.	ENTAILMENT	CONTRADICTION	UNKNOWN
18.	CONTRADICTION	CONTRADICTION	UNKNOWN
19.	ENTAILMENT	CONTRADICTION	ENTAILMENT
20.	CONTRADICTION	CONTRADICTION	CONTRADICTION
21.	ENTAILMENT	CONTRADICTION	ENTAILMENT
22.	CONTRADICTION	ENTAILMENT	CONTRADICTION
23.	ENTAILMENT	CONTRADICTION	ENTAILMENT
24.	CONTRADICTION	CONTRADICTION	CONTRADICTION
25.	ENTAILMENT	CONTRADICTION	ENTAILMENT
26.	CONTRADICTION	CONTRADICTION	CONTRADICTION
27.	ENTAILMENT	CONTRADICTION	ENTAILMENT
28.	CONTRADICTION	CONTRADICTION	CONTRADICTION
29.	ENTAILMENT	CONTRADICTION	UNKNOWN
30.	CONTRADICTION	CONTRADICTION	UNKNOWN

Dari hasil penelitian Tabel 11, menunjukkan bahwa fitur aritmatika dapat mendeteksi pasangan kalimat kontradiksi maupun *entailment* hingga 24 pasang kalimat atau tingkat akurasi hingga 80%. Sedangkan untuk sistem *Stanford RTE* tahun 2009, dapat mendeteksi pasangan kalimat *entailment* maupun kontradiksi hingga 14 pasang kalimat atau tingkat akurasinya 46,67%. Fitur aritmatika masih memiliki hasil tidak diketahui pada ID 13, 14, 17, 18, 29, dan 30 karena kata benda yang terdeteksi tidak merujuk pada kata benda yang sama dalam kalimat *Text* dan fitur aritmatika tidak dapat melakukan operasi aritmatika melebihi satu operasi aritmatika. Sedangkan sistem *Stanford RTE* tahun 2009 dapat menghasilkan deteksi yang benar dikarenakan perbandingan angka secara langsung tanpa adanya operasi aritmatika menghasilkan deteksi yang sesuai dengan hasil deteksi berdasarkan pemikiran manusia pada kolom HUMAN.

## 5. Simpulan

Kesalahan deteksi kontradiksi oleh sistem *Stanford RTE* tahun 2009 pada pasangan kalimat *Text* dan *Hypothesis* mengandung ekspresi aritmatika disebabkan oleh sistem *Stanford RTE* akan membandingkan secara langsung antara angka yang ada pada kalimat *Text* dan kalimat *Hypothesis* sehingga apabila pada kalimat *Text* membutuhkan operasi aritmatika untuk mendapatkan hasil deteksi yang benar, sistem *Stanford RTE* tahun 2009 akan mengalami kesalahan deteksi. Untuk menutupi kesalahan tersebut, maka ditambahkan fitur aritmatika dalam sistem *Stanford RTE*. Hasil analisa pada fitur aritmatika, menunjukkan bahwa ekspresi aritmatika dapat terjadi pada jenis angka yang beragam seperti numerik, tanggal, persen, mata uang, dan durasi. Namun yang dapat dilakukan oleh fitur aritmatika hanya pada jenis angka numerik, tanggal, dan persen dikarenakan tiap jenis angka diperlukan metode pendeteksian yang berbeda agar bias dilakukan operasi aritmatika. Dalam studi kasus yang digunakan, terdapat pasangan kalimat *Text* dan *Hypothesis* mengandung ekspresi aritmatika yang memerlukan operasi aritmatika lebih dari satu jenis operator aritmatika sehingga fitur aritmatika masih mengalami kesalahan deteksi.

## Daftar Pustaka

- [1] B. MacCartney, T. Grenager, M.-C. de Marneffe, D. Cer, and C. D. Manning, "Learning to recognize features of valid textual entailments", Proceedings of the main conference on

- Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics, Stroudsburg, PA, USA, pp.41–48, 2006.
- [2] M.-C. de Marneffe, A. N. Rafferty, and C. D. Manning, "Finding Contradictions in Text", *Proceeding of ACL-08*, pp.1039–1047, 2008.
- [3] S. Pado, M.-C. de Marneffe, B. MacCartney, A. N. Rafferty, E. Yeh, and C. D. Manning, "Deciding Entailment and Contradiction with Stochastic and Edit Distance-based Alignment", *TAC 2008 RTE track*, 2009.
- [4] The Stanford Natural Language Processing Group, The Stanford NLP (Natural Language Processing) Group, Stanford CoreNLP. [Online]. Available: <http://nlp.stanford.edu/software/corenlp.shtml>[Accessed: 01-Mar-2013].
- [5] T. Pedersen, S. Patwardhan, and J. Michelizzi, "WordNet::Similarity: measuring the relatedness of concepts", *Demonstration Papers at HLT-NAACL 2004*, Stroudsburg, PA, USA, pp.38–41, 2004.
- [6] D. J. E. Wall, "Arithmetic Word Problems Study Guide for McGraw-Hill's ASVAB, Education". <http://www.education.com/reference/article/introduction-asvab-arithmetic-word-problems/> [Accessed: 24-Feb-2013].
- [7] Text Analysis Conference (TAC) Past Data. <http://www.nist.gov/tac/data/past/2008/RTE-4.html> [Accessed: 24-Feb-2013].
- [8] R. Arianto, D. O. Siahaan, and A. Saikhu, "Perbaikan Metode Stanford Recognizing Textual Entailment pada Kalimat Ekspresi Aritmatika", *Prosiding Seminar Nasional Teknologi Informasi dan Multimedia*, pp.25–13 – 25–17, Jan. 2013.
- [9] Doxygen, Grammarscope, 2012.
- [10] K. Toutanova, D. Klein, C. D. Manning, and Y. Singer, "Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network", *Proceedings of HLT-NAACL 2003*, pp.252–259, 2003.
- [11] J. R. Finkel, T. Grenager, and C. D. Manning, "Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling", *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL 2005)*, pp.363–370, 2005.