

## ABSTRAK

Kurs mata uang rupiah saat ini sangat tidak stabil dan semakin menurun nilainya. Tidak stabilnya kurs akan mempengaruhi banyak hal. Untuk membantu permasalahan ini, dirancang sebuah aplikasi untuk melakukan prediksi kurs rupiah terhadap dolar Amerika dengan menganalisis berita bursa efek, berita ekonomi bisnis, dan moneter di Indonesia. Berita yang digunakan adalah 2 jenis, yaitu ringkasan berita dan berita lengkap. Pada ringkasan berita, topik yang diambil mengenai bursa efek Indonesia, sedangkan pada berita lengkap adalah ekonomi bisnis dan moneter. Untuk melakukan prediksi, digunakan salah satu metode temu balik informasi, yaitu *Latent Semantic Indexing (LSI)*. Prediksi juga dilakukan tanpa menggunakan *LSI*, tetapi dengan menggunakan kumpulan kata untuk prediksi 'naik'. Mengacu pada hasil eksperimen, disimpulkan bahwa *LSI* bukan model yang tepat untuk melakukan prediksi kurs, tetapi baik untuk mengaitkan konteks berita yang tepat, karena berita yang ada pada saat tertentu belum tentu mencerminkan naik turunnya kurs. Kemunculan kata dalam dokumen juga tidak memiliki korelasi dengan naik atau turunnya kurs karena sebuah kata belum tentu memiliki makna yang sama untuk konteks dan dokumen yang berbeda.

Kata kunci : *Latent Semantic Indexing*, prediksi, kurs mata uang, analisis berita

## **ABSTRACT**

*Currency rate fluctuates very fast in the world today. To anticipate this phenomenon, an application is designed to perform Indonesian Rupiah's rate prediction against American dollar by analyzing stock exchange, business economic and monetary news in Indonesia. Two types of news are analyzed, i.e.: summarization and full document collection. In the summarization-based collection, the topic about 'Indonesian stock exchange' is collected from Google search engine results. In the full document collection, the topics about business economic and monetary are collected from Detik Finance. To perform the prediction, an information retrieval method is utilized, i.e. the Latent Semantic Indexing (LSI). The experiment results show that LSI was not a suitable method to perform currency rate prediction. One of the reasons is that the words' co-occurrences in the documents provide some indications about the rate's movement, but not a real decision. Another reason is that the semantics of a word is not necessarily the same in various documents. Although, the LSI shows good performance for creating connection of words in news documents.*

*Keywords: Latent Semantic Indexing, prediction, currency rate, news analysis*

## DAFTAR ISI

LEMBAR PENGESAHAN .....	ii
PERNYATAAN ORISINALITAS LAPORAN PENELITIAN .....	iii
PERNYATAAN PUBLIKASI LAPORAN PENELITIAN .....	iv
PRAKATA .....	v
ABSTRAK .....	vii
ABSTRACT .....	viii
DAFTAR ISI .....	ix
DAFTAR TABEL .....	xi
DAFTAR GAMBAR .....	xii
DAFTAR GRAFIK .....	xiv
DAFTAR SIMBOL .....	xv
DAFTAR KODE PROGRAM .....	xvii
DAFTAR RUMUS .....	xviii
DAFTAR ISTILAH .....	xix
DAFTAR LAMPIRAN .....	xx
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Tujuan .....	2
1.4 Batasan Masalah .....	2
1.5 Sistematika Pembahasan .....	3
BAB II LANDASAN TEORI .....	4
2.1 Temu Balik Informasi ( <i>Information Retrieval</i> ) .....	4
2.2 <i>Latent Semantic Indexing (LSI)</i> .....	5
2.3 Java .....	9
2.4 Basis Data .....	10
2.5 <i>Entity Relationship Diagram (ERD)</i> .....	10
2.5.1 Entitas .....	10
2.5.2 Relasi .....	10
2.5.3 Atribut .....	10
2.5.4 Kardinalitas .....	11
2.6 <i>Unified Modeling Language (UML)</i> .....	13
2.6.1 <i>Uses Case Diagram</i> .....	13
2.6.2 <i>Activity Diagram</i> .....	13
2.6.3 <i>Class Diagram</i> .....	14
2.7 <i>Jsoup</i> .....	15
2.8 <i>LingPipe</i> .....	15
BAB III ANALISIS DAN DISAIN .....	16
3.1 Metodologi Penelitian .....	16
3.2 Gambaran Keseluruhan .....	18
3.2.1 Persyaratan Antarmuka Eksternal .....	18
3.2.2 Antarmuka dengan Pengguna .....	18
3.2.3 Antarmuka Perangkat Keras .....	19
3.2.4 Antarmuka Perangkat Lunak .....	19
3.2.5 Fitur-Fitur Produk Perangkat Lunak .....	19

3.3	Disain Perangkat Lunak .....	24
3.3.1	Pemodelan Perangkat Lunak.....	24
3.3.2	Disain Penyimpanan Data.....	27
3.3.3	Disain Antarmuka .....	27
BAB IV PENGEMBANGAN PERANGKAT LUNAK .....		30
4.1	Pengolahan Data.....	30
4.2	Pembentukan Matriks <i>LSI</i> .....	38
4.3	Implementasi Kelas .....	42
4.3.1	Kelas DataTanggal .....	42
4.3.2	Kelas KataTanggal .....	43
4.3.3	Kelas KamusKata.....	43
4.3.4	Kelas Stopwords .....	44
4.3.5	Kelas KursExcel.....	44
4.3.6	Kelas Berita.....	44
4.3.7	Kelas ListDataTanggal.....	45
4.3.8	Kelas ListKataTanggal.....	46
4.3.9	Kelas ListKamusKata.....	47
4.3.10	Kelas ListStopwords .....	48
4.3.11	Kelas ListKursExcel.....	49
4.3.12	Kelas ListBerita.....	49
4.3.13	Kelas AbsTableBerita .....	50
4.3.14	Kelas NewsGrabber .....	50
4.3.15	Kelas MatrixRetrieve .....	51
4.3.16	Kelas NewsTokenizer .....	51
4.3.17	Kelas TopResult.....	53
4.3.18	Kelas Lsi.....	54
4.4	Implementasi Penyimpanan Data.....	56
4.5	Implementasi Antarmuka .....	57
4.5.1	<i>Form</i> Halaman Utama.....	57
4.5.2	<i>Form</i> Berita.....	58
4.5.3	<i>Form</i> Prediksi.....	59
BAB V TESTING DAN EVALUASI SISTEM .....		61
5.1	Eksperimen.....	61
5.2	Perolehan Kumpulan Kata Untuk Prediksi ‘Naik’ .....	64
5.3	Hasil Eksperimen .....	66
5.4	Evaluasi Terhadap <i>K</i> .....	77
5.5	Evaluasi Terhadap Akurasi .....	78
5.6	Evaluasi Frekuensi Kata.....	80
5.7	Evaluasi Aplikasi.....	81
BAB VI KESIMPULAN DAN SARAN .....		85
6.1	Kesimpulan.....	85
6.2	Saran .....	85
DAFTAR PUSTAKA .....		xxi
RIWAYAT HIDUP.....		86

## DAFTAR TABEL

Tabel 2.1 Matriks <i>Term-Document C</i> (Manning, C. D., et al, 2009 : 414).....	6
Tabel 2.2 Matriks <i>U</i> (Manning, C. D., et al, 2009 : 414).....	6
Tabel 2.3 Matriks $\Sigma_k$ (Manning, C. D., et al, 2009 : 414) .....	7
Tabel 2.4 Matriks $V^T$ (Manning, C. D., et al, 2009 : 415).....	7
Tabel 2.5 Matriks $\Sigma_2$ (Manning, C. D., et al, 2009 : 415) .....	7
Tabel 2.6 Matriks $C_2$ (Manning, C. D., et al, 2009 : 415).....	8
Tabel 2.7 Jumlah Kata Pada <i>Database</i> .....	37
Tabel 4.1 Statistik Data Berita .....	32
Tabel 4.2 Statistik Jumlah Kurs .....	32
Tabel 5.1 Metode .....	61
Tabel 5.2 Skenario .....	62
Tabel 5.3 Contoh Prediksi Dengan Skenario Top-5 .....	64
Tabel 5.4 Kelebihan dan Kekurangan Metode.....	78
Tabel 5.5 Kelebihan dan Kekurangan Skenario.....	79
Tabel 5.6 <i>Test Case Form</i> Berita .....	81
Tabel 5.7 <i>Test Case Form</i> Prediksi .....	84

## DAFTAR GAMBAR

Gambar 2.1 Vektor-Vektor Dokumen Hasil Perhitungan <i>LSI</i> (Manning, C. D., et al, 2009 : 416) .....	8
Gambar 2.2 Relasi Satu ke Satu .....	12
Gambar 2.3 Relasi Satu ke Banyak .....	12
Gambar 2.4 Relasi Banyak ke Satu .....	12
Gambar 2.5 Relasi Banyak ke Banyak .....	12
Gambar 2.6 Contoh <i>Use Case Diagram</i> .....	13
Gambar 2.7 Contoh Kelas .....	14
Gambar 2.8 <i>Association</i> .....	14
Gambar 2.9 <i>Aggregation</i> .....	14
Gambar 2.10 <i>Composition</i> .....	15
Gambar 2.11 <i>Inheritance</i> .....	15
Gambar 3.1 <i>Flowchart</i> Langkah Kerja .....	16
Gambar 3.2 <i>Use Case</i> Aplikasi Prediksi Kurs Mata Uang .....	24
Gambar 3.3 <i>Use Case Sub-System</i> Mengelola Berita .....	25
Gambar 3.4 <i>Acitivity Diagram</i> Mengelola Berita .....	25
Gambar 3.5 <i>Activity Diagram</i> Prediksi .....	26
Gambar 3.6 Rancangan <i>ERD</i> .....	27
Gambar 3.7 Tampilan Rancangan <i>Form</i> Halaman Utama .....	28
Gambar 3.8 Tampilan Rancangan <i>Form</i> Berita .....	28
Gambar 3.9 Tampilan Rancangan <i>Form</i> Prediksi .....	29
Gambar 4.1 Diagram Kelas Aplikasi Prediksi Kurs .....	42
Gambar 4.2 Kelas DataTanggal .....	43
Gambar 4.3 Kelas KataTanggal .....	43
Gambar 4.4 Kelas KamusKata .....	43
Gambar 4.5 Kelas Stopwords .....	44
Gambar 4.6 Kelas KursExcel .....	44
Gambar 4.7 Kelas Berita .....	44
Gambar 4.8 Kelas ListDataTanggal .....	45
Gambar 4.9 Kelas ListKataTanggal .....	46
Gambar 4.10 Kelas ListKamusKata .....	47
Gambar 4.11 Kelas ListStopwords .....	48
Gambar 4.12 Kelas ListKursExcel .....	49
Gambar 4.13 Kelas ListBerita .....	49
Gambar 4.14 Kelas AbsTableBerita .....	50
Gambar 4.15 Kelas NewsGrabber .....	51
Gambar 4.16 Kelas MatrixRetrieve .....	51
Gambar 4.17 Kelas NewsTokenizer .....	52
Gambar 4.18 Kelas TopResult .....	53
Gambar 4.19 Kelas Lsi .....	56
Gambar 4.20 Implementasi Basis Data .....	56
Gambar 4.21 <i>Form</i> Halaman Utama .....	57
Gambar 4.22 <i>Form</i> Berita .....	58
Gambar 4.23 <i>Form</i> Prediksi .....	59

Gambar 4.24 Hasil Prediksi .....	59
Gambar 4.25 <i>Form Alasan</i> .....	60

## DAFTAR GRAFIK

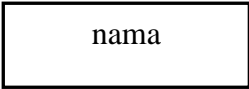
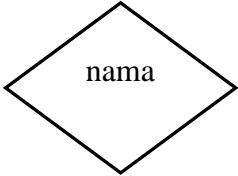




Grafik 5.1 Hasil Eksperimen <i>Term Document Matrix</i> Tanpa <i>TF</i> dengan $k = 69$ Menggunakan Ringkasan Berita .....	66
Grafik 5.2 Hasil Eksperimen <i>Term Document Matrix</i> Tanpa <i>TF</i> dengan $k = 30$ Menggunakan Ringkasan Berita .....	67
Grafik 5.3 Hasil Eksperimen <i>Term Document Matrix</i> Dengan <i>TF</i> dengan $k = 69$ Menggunakan Ringkasan Berita .....	69
Grafik 5.4 Hasil Eksperimen <i>Term Document Matrix</i> dengan <i>TF</i> dengan $k = 30$ Menggunakan Ringkasan Berita .....	70
Grafik 5.5 Metode 2 dengan Top-N Kata untuk Prediksi ‘Naik’ Menggunakan Ringkasan Berita .....	71
Grafik 5.6 Hasil Eksperimen <i>Term Document Matrix</i> Tanpa <i>TF</i> dengan $k = 69$ Menggunakan Berita Lengkap .....	72
Grafik 5.7 Hasil Eksperimen <i>Term Document Matrix</i> Tanpa <i>TF</i> dengan $k = 30$ Menggunakan Berita Lengkap .....	73
Grafik 5.8 Hasil Eksperimen <i>Term Document Matrix</i> dengan <i>TF</i> dengan $k = 69$ Menggunakan Berita Lengkap .....	74
Grafik 5.9 Hasil Eksperimen <i>Term Document Matrix</i> dengan <i>TF</i> dengan $k = 30$ Menggunakan Berita Lengkap .....	75
Grafik 5.10 Metode 2 dengan Top-N Kata untuk Prediksi ‘Naik’ Menggunakan Berita Lengkap .....	76



## DAFTAR SIMBOL


### 1. Tabel simbol *ERD*


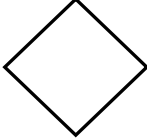
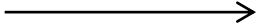

Tabel nomor 1 ini merupakan tabel mengenai simbol-simbol *ERD*. Terdapat gambar simbol dan nama dari setiap simbolnya.

Nama Simbol	Simbol
Entitas	
Relasi	
Atribut	
Atribut kunci primer	
Atribut bernilai banyak	
Atribut turunan	

### 2. Tabel Simbol *Activity Diagram*

Tabel nomor 2 merupakan tabel mengenai simbol *activity diagram*. Terdapat gambar simbol dan deskripsi untuk setiap simbolnya.

Simbol	Deskripsi
	<i>Initial state</i> untuk memulai sebuah <i>activity diagram</i> .

<b>Simbol</b>	<b>Deskripsi</b>
	<i>Action state</i> menggambarkan proses atau aktivitas pada <i>activity diagram</i>
	<i>Decision</i> untuk percabangan jika terdapat aktivitas yang memiliki kondisi tertentu untuk berhubungan dengan aktivitas lain.
	<i>Transition</i> sebagai alur penghubung antar aktivitas.
	<i>Final state</i> untuk mengakhiri sebuah <i>activity diagram</i> .

### 3. Tabel Simbol *Visibility*

Tabel nomor 3 merupakan tabel mengenai simbol *visibility* untuk diagram kelas. Terdapat bentuk simbol dan deskripsi untuk setiap simbolnya.

<b>Simbol</b>	<b>Deskripsi</b>
+	<i>Public visibility</i> berfungsi agar atribut atau operasi dapat diakses oleh semua kelas.
#	<i>Protected visibility</i> berfungsi agar atribut atau operasi hanya dapat diakses oleh turunan dari kelas tersebut.
~	<i>Package Visibility</i> berfungsi agar atribut atau operasi hanya dapat diakses oleh kelas pada <i>package</i> yang sama.
-	<i>Private Visibility</i> berfungsi agar atribut atau operasi hanya dapat diakses kelas itu sendiri.

## DAFTAR KODE PROGRAM

Program 4.1 Potongan Kode Pengambilan Ringkasan Berita Pada <i>Google</i> .....	32
Program 4.2 Potongan Kode Pengambilan Ringkasan Berita Pada Detik .....	33
Program 4.3 Potongan Kode Tokenisasi Ringkasan Berita Untuk <i>Database</i> .....	36
Program 4.4 Potongan Kode Pembentukan <i>Term Document Matrix</i> Tanpa <i>TF</i> ...	38
Program 4.5 Potongan Kode Pembentukan <i>Term Document Matrix</i> dengan <i>TF</i> .	39
Program 4.6 Potongan Kode Pembentukan 3 Buah Matriks.....	41
Program 5.1 Potongan Kode Pembentukan <i>Query Vector</i> .....	63
Program 5.2 Potongan Kode <i>Cosine Similarity</i> .....	63
Program 5.3 Potongan Kode Perolehan Kumpulan Kata untuk Prediksi ‘Naik’ ..	65

## DAFTAR RUMUS

Rumus 2.1 Menghitung <i>Term Document Matrix</i> Dari Hasil Perhitungan <i>SVD</i> .....	5
Rumus 2.2 Menghitung Vektor <i>Query LSI</i> .....	8
Rumus 2.3 Menghitung <i>Cosine Similarity</i> Antara <i>Query</i> Dengan Dokumen .....	9

## DAFTAR ISTILAH

No	Istilah	Deskripsi
1	<i>Library</i>	Koleksi dari rutin-rutin program yang digunakan untuk membangun dan mengembangkan perangkat lunak (Wikipedia, 2014).
2	<i>Software</i>	Data yang diformat dan disimpan secara digital, termasuk program komputer, dokumentasinya, dan berbagai informasi yang bisa dibaca dan ditulis oleh komputer (Wikipedia, 2014).
3	<i>URL (Uniform Resource Locator)</i>	Alamat yang mengidentifikasi <i>file</i> tertentu pada Internet, biasanya terdiri dari protokol seperti http, diikuti oleh nama domain (Dictionary.com, 2014).
4	<i>LSI (Latent Semantic Indexing)</i>	<i>LSI</i> merupakan sebuah metode <i>indexing</i> yang memasukkan <i>queries</i> ke dalam representasi tingkat rendah, dan juga memungkinkan untuk menghitung nilai kemiripan dari <i>query</i> dokumen ke dalam representasi tingkat rendah dari matriks <i>term-document</i> menggunakan <i>Singular Value Decomposition (SVD)</i> (Manu, 2006; Manning, Raghavan, Schutze, 2009).

## DAFTAR LAMPIRAN

LAMPIRAN A .....	A-1
LAMPIRAN B .....	B-1