

## DAFTAR PUSTAKA

- Andriani, Mirna dan Manurung, Ruli. 2009. *A Survey of Bahasa Indonesia NLP Reasearch Conducted at The University of Indonesia*. Indonesia: Faculty of Computer Science University of Indonesia
- Baoli, Li., Shiwen, Yu., dan Qin, Lu. 2003. *An Improved k-Nearest Neighbors for Text Categorization*. China: To appear in the Proceedings of the 20<sup>th</sup> International Conference of Computer Processing of Oriental Language.
- Buss, Keno. 2008. *Literature Review on Preprocessing for Text Mining*. UK: De Montfort University
- Chairina, Putri., Rizal, Achmad., dan Agung B. W, Tjokorda. 2008. *Pengenalan Individu berdasarkan GAIT menggunakan PCA dan k-Nearest Neighbor*. Bandung: Departemen Teknik Informatika Institut Teknologi Telkom.
- Christiana, Priskilla. 2009. *Penerapan Metode Improved KNN Berbasis Clustering untuk Pengkategorian Dokumen Berita Berbahasa Indonesia*. Malang: Universitas Brawijaya
- Daqiqil, Ibnu. 2008. *Perancangan Dan Penerapan Algoritma Stemming Kata Berbahasa Indonesia*. Malang : Universitas Brawijaya
- Destuardi dan Sumpeno Surya. 2008. *Klasifikasi Emosi Untuk Teks Bahasa Indonesia Menggunakan Metode Naïve Bayes*. Surabaya : Institut Teknologi Sepuluh Nopember
- Even, Yair dan Zohar. 2002. *Itroudction to Text Mining*. Illionis: University Of Illionis
- Fan, Weiguo., et al. 2005. *Tapping The Power of Text Mining*. (<http://filebox.vt.edu>)
- Guo, dkk . 2002. *An kNN Model-based Approach and Its Application in Text Categorization*. UK: School of Computing and Mathematics, University of Ulster Newtownabbey
- Hearst, Marti. 2003. *What is Text Mining?*. SIMS UC Berkeley. (Online). (<http://people.ischool.berkeley.edu/~hearst/text-mining.html>, diakses 30 September 2010)
- Hermawandi, Dudi. 2008. *Implementasi Pembobotan SICBI pada Aplikasi Essay Grading Metode Latent Semantic Analisis*. Indonesia: Universitas Indonesia

- Larose, T. , Daniel. 2005. *Discovering Knowledge In Data: An Introduction To Data Mining*. New Jersey : John Wiley & Sons, Inc
- Mandala, Rila. 2006. *Evaluasi Kinerja Sistem Penyaringan Informasi Model Ruang Vektor*. Bandung: STEI Intitut Teknologi Bandung
- Mooney, Raymond J. 2006. *Machine Learning Text Categorization*. Austin: University of Texas
- Nurdiana, W. 2011. *Penerapan Algoritma Improved K-Nearest Neighbors Untuk Pengkategorian Dokumen Berita Berbahasa Indonesia*. Malang : Universitas Brawijaya Malang
- Salton, G., Wong A., dan Yang C.S. 1975. *A Vector Space Model for Automatic Indexing*. Cornell University
- Sebastiani, Fabrizio. 2002. *Text Categorization*. Italy: Dipartimentodi Matematica Purae Applicata Universitadi Padova
- Singhal, Amitabh Kumar. 1997. *Term Weighting Revisited*. Cornell University
- Z. Tala, Fadillah. 2003. *A Study of Stemming Effect On Information Retrieval In Bahasa Indonesia*. Netherlands: Institute for Logic Language and Computation Universiteit van Amsterdam
- Yang, Yiming dan Joachims, Thorsten.2008. *Text Categorization*. USA: Scholarpedia



**Lampiran**  
**Daftar *Stopwords***

No	Kata	No	Kata	No	Kata
1	ada	34	sebagai	67	sebetulnya
2	adanya	35	sebagainya	68	biasa
3	adalah	36	bagaimana	69	biasanya
4	adapun	37	bagaimanapun	70	bila
5	agak	38	sebagaimana	71	bilakah
6	agaknya	39	bagaimanakah	72	bisa
7	agar	40	bagi	73	bisakah
8	akan	41	bahkan	74	sebisanya
9	akankah	42	bahwa	75	boleh
10	akhirnya	43	bahwasannya	76	bolehkah
11	aku	44	sebaliknya	77	bolehlah
12	akulah	45	banyak	78	buat
13	amat	46	sebanyak	79	bukan
14	amatlah	47	beberapa	80	bukankah
15	anda	48	seberapa	81	bukanlah
16	andalah	49	begini	82	bukannya
17	antar	50	beginilah	83	cuma
18	diantaranya	51	sebegini	84	percuma
19	antara	52	begitu	85	dahulu
20	antaranya	53	begitukah	86	dalam
21	diantara	54	begitulah	87	dan
22	apa	55	bagitupun	88	dapat
23	apaan	56	sebegitu	89	dari
24	mengapa	57	belum	90	daripada
25	apabila	58	belumah	91	dekat
26	apakah	59	sebelum	92	demi
27	apalagi	60	sebelumnya	93	demikian
28	apatah	61	sebenarnya	94	demikianlah
29	atau	62	berapa	95	sedemikian
30	ataukah	63	berapakah	96	dengan
31	ataupun	64	berapalah	97	depan
32	bagai	65	berapapun	98	di
33	bagaimana	66	betulkah	99	dia

No	Kata	No	Kata	No	Kata
100	dialah	136	itulah	172	kinilah
101	dini	137	jangan	173	kiranya
102	diri	138	janganakan	174	sekiranya
103	dirinya	139	janganlah	175	kita
104	terdiri	140	jika	176	kitalah
105	dong	141	jikalau	177	kok
106	dulu	142	juga	178	lagi
107	enggak	143	justru	179	lagian
108	enggaknya	144	kalau	180	lah
109	entah	145	kalaulah	181	lain
110	entahlah	146	kalaupun	182	lainnya
111	terhadap	147	berkali	183	melainkan
112	terhadapnya	148	kali	184	selaku
113	hal	149	sekali	185	lalu
114	hampir	150	kalian	186	melalui
115	hanya	151	kami	187	terlalu
116	hanyalah	152	kamilah	188	lama
117	harus	153	kamu	189	lamanya
118	haruslah	154	kamulah	190	selama
119	harusnya	155	kan	191	selamanya
120	hendak	156	kapan	192	lebih
121	hendaklah	157	kapankah	193	terlebih
122	hendaknya	158	kapanpun	194	bermacam
123	hingga	159	dikarenakan	195	macam
124	sehingga	160	karena	196	semacam
125	ia	161	karenanya	197	maka
126	ialah	162	ke	198	makanya
127	ibarat	163	kecil	199	makin
128	ingin	164	kemudian	200	malah
129	inginkah	165	kenapa	201	malahan
130	inginkan	166	kepada	202	mampu
131	ini	167	kepadanya	203	mampukah
132	inikah	168	ketika	204	mana
133	inilah	169	seketika	205	manakala
134	itu	170	khususnya	206	manalagi
135	itukah	171	kini	207	masih

No	Kata	No	Kata	No	Kata
208	masihkah	244	pula	280	sejenak
209	semasih	245	pun	281	sekali
210	masing	246	merupakan	282	sekalian
211	mau	247	rupanya	283	sekalipun
212	maupun	248	serupa	284	sesekali
213	semaunya	249	saat	285	sekaligus
214	memang	250	saatnya	286	sekarang
215	mereka	251	sesaat	287	sekitar
216	merekalah	252	saja	288	sekitarnya
217	meski	253	sajalah	289	sela
218	meskipun	254	saling	290	selain
219	semula	255	bersama	291	seluruh
220	mungkin	256	sama	292	seluruhnya
221	mungkinkah	257	sesama	293	semakin
222	nah	258	sambil	294	sementara
223	namun	259	sampai	295	sempat
224	nanti	260	sana	296	semua
225	nantinya	261	sangat	297	semuanya
226	nyaris	262	sangatlah	298	sendiri
227	oleh	263	saya	299	sendirinya
228	olehnya	264	sayalah	300	seolah
229	seorang	265	se	301	seperti
230	seseorang	266	sebab	302	sepertinya
231	pada	267	sebabnya	303	sering
232	padanya	268	sebuah	304	seringnya
233	padahal	269	tersebut	305	serta
234	paling	270	tersebutlah	306	siapa
235	sepanjang	271	sedang	307	siapakah
236	pantas	272	sedangkan	308	siapapun
237	sepantasnya	273	sedikit	309	disini
238	sepantasnyalah	274	sedikitnya	310	disinilah
239	para	275	segala	311	sini
240	pasti	276	segalanya	312	sinilah
241	pastilah	277	segera	313	sesuatu
242	per	278	sesegera	314	sesuatunya
243	pernah	279	sejak	315	suatu



No	Kata	No	Kata	No	Kata
316	sesudah	329	tentu	342	tidaklah
317	sesudahnya	330	tentulah	343	toh
318	sudah	331	tentunya	344	waduh
319	sudahkah	332	tertentu	345	wah
320	sudahlah	333	seterusnya	346	wahai
321	supaya	334	tapi	347	sewaktu
322	tadi	335	tetapi	348	walau
323	tadinya	336	setiap	349	walaupun
324	tak	337	tiap	350	wong
325	tanpa	338	setidak	351	yaitu
326	setelah	339	setidaknya	352	yakni
327	telah	340	tidak	353	yang
328	tentang	341	tidakkah		

*Suggested Stoplist*

No	Kata	No	Kata	No	Kata
354	berada	374	bagian	394	diberi
355	keadaan	375	sebagian	395	diberikan
356	akhir	376	baik	396	memberi
357	akhiri	377	sebaik	397	memberikan
358	berakhir	378	baiknya	398	besar
359	berakhirilah	379	sebaiknya	399	sebesar
360	berakhirnya	380	bakal	400	betul
361	diakhiri	381	bakalan	401	kebetulan
362	diakhirinya	382	balik	402	dibuat
363	mengakhiri	383	terbanyak	403	dibuatnya
364	terakhir	384	bapak	404	diperbuat
365	artinya	385	baru	405	diperbuatnya
366	berarti	386	bawah	406	membuat
367	asal	387	belakang	407	memperbuat
368	asalkan	388	belakangan	408	bulan
369	atas	389	benar	409	bung
370	awal	390	benarkah	410	cara
371	awalnya	391	benarlah	411	caranya
372	berawal	392	beri	412	secara
373	berbagai	393	berikan	413	cukup

No	Kata	No	Kata	No	Kata
414	cukupkah	450	ikut	486	sejumlah
415	cukuplah	451	diingat	487	sekadar
416	secukupnya	452	diingatkan	488	sekadarnya
417	terdahulu	453	ingat	489	berkata
418	didapat	454	mengingat	490	dikatakan
419	mendapat	455	mengingatkan	491	dikatakannya
420	mendapatkan	456	seingat	492	kata
421	terdapat	457	teringat	493	katakan
422	berdatangan	458	berkeinginan	494	katakanlah
423	datang	459	diinginkan	495	katanya
424	didatangkan	460	keinginan	496	mengatakan
425	mendatang	461	menginginkan	497	mengatakannya
426	mendatangi	462	jadi	498	sekecil
427	mendatangkan	463	jadilah	499	keluar
428	dua	464	jadinya	500	kembali
429	kedua	465	menjadi	501	berkenan
430	keduanya	466	terjadi	502	mengenai
431	empat	467	terjadilah	503	bekerja
432	seenaknya	468	terjadinya	504	dikerjakan
433	digunakan	469	jauh	505	mengerjakan
434	dipergunakan	470	sejauh	506	dikira
435	gunakan	471	dijawab	507	diperkirakan
436	gunakan	472	jawab	508	kira
437	mempergunakan	473	jawaban	509	memperkirakan
438	menggunakan	474	jawabnya	510	mengira
439	hari	475	menjawab	511	terkira
440	berkehendak	476	dijelaskan	512	kurang
441	menghendaki	477	dijelaskannya	513	sekurang
442	diibaratkan	478	jelas	514	kurangnya
443	diibaratkannya	479	jelaskan	515	sekurangnya
444	ibaratkan	480	jelaslah	516	berlainan
445	ibaratnya	481	jelasnya	517	dilakukan
446	mengibaratkan	482	menjelaskan	518	melakukan
447	ibu	483	berjumlah	519	berlalu
448	berikut	484	jumlah	520	dilalui
449	berikutnya	485	jumlahnya	521	keterlalu

No	Kata	No	Kata	No	Kata
522	kelamaan	558	mirip	594	memerlukan
523	berlangsung	559	dimisalkan	595	perlu
524	lanjut	560	memisalkan	596	perlukah
525	lanjutnya	561	misal	597	perlunya
526	selanjutnya	562	misalkan	598	seperlunya
527	berlebihan	563	misalnya	599	pertama
528	lewat	564	semisal	600	tama
529	dilihat	565	semisalnya	601	memihak
530	diperlihatkan	566	bermula	602	pihak
531	kelihatan	567	mulanya	603	pihaknya
532	kelihatannya	568	dimulai	604	sepihak
533	melihat	569	dimulailah	605	pukul
534	melihatnya	570	dimulainya	606	dipunyai
535	memperlihatkan	571	memulai	607	mempunyai
536	terlihat	572	mulai	608	punya
537	kelima	573	mulailah	609	merasa
538	lima	574	dimungkinkan	610	rasa
539	luar	575	kemungkinan	611	terasa
540	bermaksud	576	kemungkinannya	612	rata
541	dimaksud	577	memungkinkan	613	berupa
542	dimaksudkan	578	menaiki	614	disampaikan
543	dimaksudkannya	579	naik	615	kesampaian
544	dimaksudnya	580	menanti	616	menyampaikan
545	semampu	581	menantikan	617	sampaikan
546	semampunya	582	nanti	618	sampai
547	masa	583	menyatakan	619	tersampaikan
548	semasa	584	nyatanya	620	menyangkut
549	masalah	585	ternyata	621	satu
550	masalahnya	586	pak	622	disebut
551	termasuk	587	panjang	623	disebutkan
552	semata	588	dipastikan	624	disebutkannya
553	diminta	589	memastikan	625	menyebutkan
554	dimintai	590	penting	626	sebut
555	meminta	591	pentingnya	627	sebutlah
556	memintakan	592	diperlukan	628	sebutnya
557	minta	593	diperlukannya	629	keseluruhan



No	Kata	No	Kata	No	Kata
630	keseluruhannya	665	menanyakan	700	tunjuk
631	menyeluruh	666	pertanyaan	701	berturut
632	sendirian	667	pertanyakan	702	menurut
633	bersiap	668	tanya	703	turut
634	mempersiapkan	669	tanyakan	704	berturut
635	menyiapkan	670	tanyanya	705	dituturkan
636	siap	671	ditegaskan	706	dituturkannya
637	dipersoalkan	672	menegaskan	707	menuturkan
638	mempersoalkan	673	tegas	708	tutur
639	persoalan	674	tegasnya	709	tuturnya
640	soal	675	setempat	710	diucapkan
641	soalnya	676	tempat	711	diucapkannya
642	diketahui	677	setengah	712	mengucapkan
643	diketahui	678	tengah	713	mengucapkannya
644	mengetahui	679	tepat	714	ucap
645	tahu	680	terus	715	ucapnya
646	tahun	681	tetap	716	berujar
647	ditambahkan	682	setiba	717	ujar
648	menambahkan	683	setibanya	718	ujarnya
649	tambah	684	tiba	719	umum
650	tambahnya	685	tiga	720	umumnya
651	tampak	686	setinggi	721	diungkapkan
652	tampaknya	687	tinggi	722	mengungkapkan
653	ditandaskan	688	ditujukan	723	ungkap
654	menandaskan	689	menuju	724	ungkapnya
655	tandas	690	terturu	725	untuk
656	tandasnya	691	ditunjuk	726	usah
657	bertanya	692	ditunjuki	727	seusai
658	dipertanyakan	693	ditunjukkan	728	usai
659	ditanya	694	ditunjukkannya	729	terutama
660	ditanyai	695	ditunjuknya	730	waktu
661	ditanyakan	696	menunjuk	731	waktunya
662	mempertanyakan	697	menunjuki	732	meyakini
663	menanya	698	menunjukkan	733	meyakinkan
664	menanyai	699	menunjuknya	734	yakin