

Summary Percobaan Mencari Parameter Multi Rake Yang Optimal

Daftar Isi

Pendahuluan.....	1
Perbandingan Hasil.....	2
Teks 1.....	2
Hasil Code 1.....	2
Kesimpulan Code 1.....	3
Hasil Code 2.....	3
Kesimpulan Code 2.....	3
Teks 2.....	3
Hasil Code 1.....	4
Kesimpulan Code 1.....	4
Hasil Code 2.....	4
Kesimpulan Code 2.....	4
Teks 3.....	4
Hasil Code 1.....	5
Kesimpulan Code 1.....	5
Hasil Code 2.....	5
Kesimpulan Code 2.....	5
Teks 4.....	6
Hasil Code 1.....	6
Kesimpulan Code 1.....	6
Hasil Code 2.....	7
Kesimpulan Code 2.....	7
Kesimpulan.....	7
Perbandingan Hasil Beberapa Percobaan Parameter.....	8
Parameter 1.....	8
Parameter 2.....	8
Parameter 3.....	9
Parameter 4.....	9
Parameter 5.....	10
Parameter 6.....	10
Parameter 7.....	11
Kesimpulan Perbandingan Hasil Beberapa Percobaan Parameter.....	11

Pendahuluan

Pada code sebelumnya paramater Multi Rake sebagai berikut.

```
rake = Rake(  
    min_chars=3,  
    max_words=3,  
    min_freq=1,  
    language_code='id', # 'en'  
    stopwords=stopwordsid2, # {'and', 'of'}  
    lang_detect_threshold=10,  
    max_words_unknown_lang=2,  
    generated_stopwords_percentile=10,  
    generated_stopwords_max_len=3,  
    generated_stopwords_min_freq=1, )
```

Penjelasan:

1. **min_chars=3**: Jumlah karakter minimum yang diperlukan agar sebuah kata dianggap sebagai kata kunci. Dalam hal ini, setiap kata kunci harus memiliki setidaknya 3 karakter.
2. **max_words=3**: Jumlah maksimum kata dalam sebuah frasa yang dianggap sebagai kata kunci. Hanya tiga kata pertama dari setiap frasa yang dianggap sebagai kata kunci.
3. **min_freq=1**: Frekuensi minimum yang diperlukan agar sebuah kata dianggap sebagai kata kunci. Kata yang muncul sekali atau lebih dianggap sebagai kata kunci.
4. **language_code='id'**: Kode bahasa yang digunakan untuk pemrosesan teks. Dalam hal ini, bahasa yang digunakan adalah Bahasa Indonesia ('id').
5. **stopwords=stopwordsid2**: Daftar kata-kata yang dianggap sebagai stopwords (kata-kata umum yang diabaikan dalam pemrosesan teks). Daftar ini didefinisikan oleh variabel stopwordsid2.
6. **lang_detect_threshold=10**: Ambang batas untuk mendeteksi bahasa dalam teks. Jika skor deteksi bahasa lebih besar dari 10, teks dianggap sebagai bahasa yang ditentukan; sebaliknya, teks dianggap sebagai teks campuran.
7. **max_words_unknown_lang=2**: Jumlah kata maksimum yang dipertahankan saat menangani teks dengan bahasa yang tidak diketahui.
8. **generated_stopwords_percentile=10**: Persentil dari distribusi frekuensi kata yang dihasilkan dan ditambahkan sebagai stopwords. Persentil 10% terendah dari frekuensi.
9. **generated_stopwords_max_len=3**: Panjang maksimum kata yang dihasilkan yang dianggap sebagai stopwords. Kata-kata yang dihasilkan dengan panjang lebih dari 3 karakter dianggap sebagai stopwords.
10. **generated_stopwords_min_freq=1**: Frekuensi minimum yang diperlukan agar sebuah kata yang dihasilkan dianggap sebagai stopwords. Kata yang dihasilkan dengan frekuensi satu kali atau lebih dianggap sebagai stopwords.

Diketahui dengan parameter tersebut hasil text yang relevan atau yang bisa menjadi suatu ringkasan dalam suatu paragraf/kalimat sangat sedikit. Hasil bisa dilihat [disini](#)

Lalu setelah dilakukan beberapa percobaan mengganti parameter, didapatkan hasil yang cukup optimal dengan mengganti nilai beberapa parameternya sebagai berikut.

```
rake = Rake(
    min_chars=20,
    max_words=50,
    min_freq=1,
    language_code='id', # 'en'
    stopwords=stopwordsid2, # {'and', 'of'}
    lang_detect_threshold=10,
    max_words_unknown_lang=3,
    generated_stopwords_percentile=10,
    generated_stopwords_max_len=3,
    generated_stopwords_min_freq=1, )
```

Pada code diatas, nilai parameter-parameter yang diganti yaitu, *min_chars=5*, *max_words=10*, dan *max_words_unknown_lang=3*.

Parameter-paramater tersebut sudah bisa dibilang cukup optimal setelah membandingkan dari hasil parameter sebelumnya.

Perbandingan Hasil

Teks 1

"\n\n loading...\n\n IHSG sepekan sepekan naik 1,09 persen. Foto/Dok\n JAKARTA - Selama pekan 28 Maret-1 April 2022, indeks harga saham gabungan (IHSG) menguat disusul dengan kapitalisasi pasar di Bursa Efek Indonesia (BEI) yang meningkat dibanding pekan lalu.\n\n \"kenaikan IHSG sebesar 1,09% menjadi 7.078 dari posisi 7.002 pada pekan sebelumnya,\" ujar Sekretaris Perusahaan BEI, Aulia Noviana Utami Putri, dalam keterangan tertulis, dikutip Minggu (3/4/2022).\n\n Senada, kenaikan juga terjadi pada kapitalisasi pasar bursa yang ditutup meningkat 1,10% menjadi Rp8.909,4 triliun dari Rp8.812,3 triliun pada pekan sebelumnya. Namun, terjadi penurunan pada rata-rata nilai transaksi harian (RNTH) sebesar 0,97% menjadi Rp13,839 triliun dari Rp13,975 triliun pada pekan yang lalu.\n\n Rata-rata volume transaksi harian bursa selama sepekan juga terjadi penurunan sebesar 10,35% menjadi 22,536 miliar saham dari 25,137 miliar saham pada penutupan pekan sebelumnya.\n\n \"Data rata-rata frekuensi harian bursa selama sepekan turut mengalami perubahan (penurunan) sebesar 3,14% menjadi 1.242.024 transaksi dari 1.282.303 transaksi pada pekan sebelumnya,\" kata dia.\n\n Pada hari perdagangan akhir pekan, investor asing mencatatkan nilai beli bersih sebesar Rp1,2 triliun dan sepanjang tahun 2022 investor asing mencatatkan beli bersih sebesar Rp33,3 triliun.\n\n Total emisi obligasi dan sukuk yang sudah tercatat sepanjang tahun 2022 adalah 33 emisi dari 27 emiten senilai Rp38,74 triliun. Dengan pencatatan ini maka total emisi obligasi dan sukuk yang tercatat di BEI berjumlah 491 emisi dengan nilai nominal outstanding sebesar Rp449,60 triliun dan USD47,5 juta, diterbitkan oleh 125 emiten.\n\n Surat berharga negara (SBN) tercatat di BEI berjumlah 151 seri dengan nilai nominal Rp4.833 triliun dan USD200,65 juta. \n\n \n (uka)\n\n \n"

Hasil Code 1

```
['investor asing mencatatkan', 'surat berharga negara', 'data rata-rata frekuensi']
```

Kesimpulan Code 1

Kesimpulan dari 3 keyword teratas sebagai berikut.

- 'investor asing mencatatkan' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)
- 'surat berharga negara' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)
- 'data rata-rata frekuensi' = Keywords ini belum relevan terhadap teks dikarenakan tidak menjelaskan apa-apa atau belum cukup mewakili ringkasan awal pada teks (*Tidak Relevan*)

Hasil Code 2

['investor asing mencatatkan beli bersih', 'aulia noviana utami putri', 'indeks harga saham gabungan']

Kesimpulan Code 2

Kesimpulan dari 3 keyword teratas sebagai berikut.

- 'investor asing mencatatkan beli bersih' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)
- 'aulia noviana utami putri' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal keunikan dan membuat penasaran dari pembaca terhadap teks oleh karena itu merupakan poin penting dari teks tersebut. (*Relevan*)
- 'indeks harga saham gabungan' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)

Teks 2

"\n \n \n \n \n \n \n \n \n \n \n \n \n Selasa, 05 April 2022 - 07:34 WIB\n \n \n \n \n loading...\n \n \n \n \n \n \n IHSG hari ini berpotensi menguat terbatas. Arsip FotoVMPIV/Faisal Rahman\n \n \n A A A \n \n JAKARTA - Indeks Harga Saham Gabungan (IHSG) hari ini berpotensi menguat terbatas pada sepanjang perdagangan. Indeks diproyeksikan bergerak di rentang 6.988 - 7.123.\n \n \n CEO Indosurya Bersinar Sekuritas William Surya Wijaya mengatakan, pergerakan IHSG kembali mencetak rekor tertinggi atau all time high (ATH) baik secara intraday maupun penutupannya.\n \n \n "Kenaikan yang terjadi dalam pola gerak IHSG saat ini masih ditopang oleh arus deras capital inflow yang masih terjadi secara ytd (year to date)," kata William dalam risetnya, Selasa (5V4V2022).\n \n \n Namun, lanjut dia, kenaikan yang terjadi terlihat sudah cukup terbatas sehingga adanya risiko koreksi wajar masih perlu diwaspadai oleh para investor.\n \n \n William lantas membeberkan sejumlah saham yang dapat dijadikan pilihan pada hari ini, di antaranya:\n \n - BBN\n \n - EXCL\n \n - ASII\n \n - PWON\n \n - SMGR\n \n - AAL\n \n - ICBP\n \n \n \n Sebelumnya, IHSG kemarin mendarat di zona hijau pada sesi terakhir perdagangan di awal pekan. Indeks ditutup naik 37,45 poin atau 0,53% tembus rekor di 7.116.\n \n \n

Terdapat 274 saham menguat, 240 saham melemah dan 175 saham stagnan. Transaksi perdagangan mencapai Rp11,2 triliun dari 20,8 miliar lembar saham yang diperdagangkan. (ind) 10 menit yang lalu 2 jam yang lalu 3 jam yang lalu 3 jam yang lalu 8 jam yang lalu 9 jam yang lalu 10 jam yang lalu 10 jam yang lalu 11 jam yang lalu Part of MNC Portal Indonesia

Hasil Code 1

['mencetak rekor tertinggi', 'all time high', 'risiko koreksi wajar']

Kesimpulan Code 1

Kesimpulan dari 3 keyword teratas sebagai berikut.

- 'mencetak rekor tertinggi' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)
- 'all time high' = Keywords ini belum relevan terhadap teks dikarenakan tidak menjelaskan apa-apa atau belum cukup mewakili ringkasan awal pada teks (*Tidak Relevan*)
- 'risiko koreksi wajar' = Keywords ini belum relevan terhadap teks dikarenakan tidak menjelaskan apa-apa atau belum cukup mewakili ringkasan awal pada teks (*Tidak Relevan*)

Hasil Code 2

['ceo indosurya bersinar sekuritas william surya wijaya mengatakan', 'arsip foto/mpi/faisal rahman a a jakarta', 'arus deras capital inflow']

Kesimpulan Code 2

Kesimpulan dari 3 keyword teratas sebagai berikut.

- 'ceo indosurya bersinar sekuritas william surya wijaya mengatakan' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)
- 'arsip foto/mpi/faisal rahman a a jakarta' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)
- 'arus deras capital inflow' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)

Teks 3

"Adi Haryanto, Koran SI 00b7 Senin 04 April 2022 10:32 WIB BANDUNG BARAT - Dua balita beradik di Kampung Cibeunying, RT 03/RW 13, Desa Cibodas, Kecamatan Lembang, Kabupaten Bandung Barat

(KBB), tenggelam di kolam renang vila yang ada di kawasan tersebut.\n\"Betul ada kejadian itu. Korbannya bernama Riri (3) dan Rafa (4) yang tewas tenggelam di kolam renang,\" kata Kepala Desa Cibodas Dindin Sukaya saat dikonfirmasi, Senin (4V4V2022).\nKejadian nahas tersebut baru diketahui pada Sabtu (2V4V2022) sekitar pukul 17.00 WIB. Balita yang merupakan kakak beradik itu tewas tenggelam di kolam renang villa Bellevue yang kedalamannya sekitar 1-1,5 meter.\nIa menuturkan, keduanya masuk ke area villa tanpa diketahui orangtua korban dan penjaga villa. Mereka diduga kuat main di dekat kolam renang villa dan tercebur, sampai ditemukan mengambang dan sudah meninggal dunia.\nSementara penjaga villa diketahui sedang salat ashar. Namun, karena posisi villa dan permukiman yang berdekatan serta pintu gerbang yang tidak dikunci, membuat kedua anak tersebut bisa masuk ke villa.\n\"Memang sepertinya tidak ada yang melihat karena penjaga villa juga sedang menunaikan salat, jadi luput dari perhatian,\" tuturnya.\n\nMenurutnya, kedua orang tua korban merasa terpuak dengan kejadian tersebut. Namun, mereka sudah merelakan dan menerima kejadian ini sebagai musibah. Kedua balita kakak beradik tersebut juga sudah dimakamkan pihak keluarga.\n\"Orang tuanya menganggap kejadian ini sebuah musibah dan persoalan ini juga sudah diselesaikan secara musyawarah dan kekeluargaan,\" tuturnya.\n

Hasil Code 1

['koran si .', 'orang tua korban', 'orang tuanya menganggap']

Kesimpulan Code 1

Kesimpulan dari 3 keyword teratas sebagai berikut.

- 'koran si .' = Keywords ini belum relevan terhadap teks dikarenakan tidak menjelaskan apa-apa atau belum cukup mewakili ringkasan awal pada teks (*Tidak Relevan*)
- 'orang tua korban' = Keywords ini belum relevan terhadap teks dikarenakan tidak menjelaskan apa-apa atau belum cukup mewakili ringkasan awal pada teks (*Tidak Relevan*)
- 'orang tuanya menganggap' = Keywords ini belum relevan terhadap teks dikarenakan tidak menjelaskan apa-apa atau belum cukup mewakili ringkasan awal pada teks (*Tidak Relevan*)

Hasil Code 2

['kepala desa cibodas dindin sukaya', 'kolam renang vila bellevue', 'koran si .']

Kesimpulan Code 2

Kesimpulan dari 3 keyword teratas sebagai berikut.

- 'kepala desa cibodas dindin sukaya' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)
- 'kolam renang vila bellevue' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)

- 'kabupaten bandung barat' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)

Teks 4

"\n\nIsmail Sangaji, iNews \u00b7 Minggu 03 April 2022 04:54 WIB\nTERNATE - Kecelakaan beruntun melibatkan tiga unit mobil dump truck terjadi di persimpangan jalan antara Smelter L dan Smelter K perusahaan PT IWIP, di Desa Lelilef, Weda Tengah, Kabupaten Halmahera Tengah, Provinsi Maluku Utara Sabtu pagi, (2/4/2022),\nAkibatnya, 11 orang karyawan dilarikan ke Klinik PT.IWIP di Tanjung Uli. Dari jumlah korban dalam kecelakaan itu terdapat 5 orang mengalami luka parah. Mereka direncanakan akan dirujuk ke RSUD Chasan Boesoirie di Kota Ternate.\nInformasi yang diterima, kronologis kejadian ini berawal dari satu unit mobil truck DT jenis shacman warna merah dengan nomor lambung 493 bermuatan Ore Nikel. Mobil yang dikendarai oleh Sulfani Alfaeni itu melaju dari arah Utara ke Selatan, namun saat di persimpangan jalan tersebut rem diduga blong sehingga mobil kehilangan kendali.\nAkibatnya mobil truck DT tersebut menabrak mobil DT lainnya dengan nomor lambung 386 dan satu unit mobil Distro dengan nomor lambung 24 yang bermuatan penumpang.\nDari jumlah korban dalam insiden ini terdapat 7 orang TKA asal China. Mereka dengan nama Jia Wei Jiang, Song Qing, Ma Yong Kang, Liu Bau Qun, Jin Chao, Wu Jin Jieo, dan Li Guo Lih. Untuk korban WNI yang merupakan tenaga kerja lokal berjumlah 4 orang diantaranya La Ode Ridwan, M Yusup, La Ode Usman, dan Sulfani Alfaeni.\nBilal Sau, Supervisor Komunikasi PT Indonesia Weda Bay Industrial Park atau IWIP, ketika dikonfirmasi, mengatakan bahwa pihak PT IWIP saat ini sedang melakukan investigasi.\nKasat Reskrim Polres Halmahera Tengah IPTU Taufik Saimima mengatakan bahwa pihak kepolisian belum mendapatkan laporan kecelakaan tersebut.\n\u201cBelum ada laporan ke kami,\u201d jelasnya."

Hasil Code 1

['rsud chasan boesoirie', 'ma yong kang', 'liu bau qun']

Kesimpulan Code 1

Kesimpulan dari 3 keyword teratas sebagai berikut.

- 'rsud chasan boesoirie' = Keywords ini belum relevan terhadap teks dikarenakan tidak menjelaskan apa-apa atau belum cukup mewakili ringkasan awal pada teks (*Tidak Relevan*)
- 'ma yong kang' = Keywords ini relevan dengan teksnya dikarenakan bisa keyword tersebut merupakan nama korban dari kecelakaan, jadi menurut saya cukup bisa menjadi ringkasan awal terhadap teks. (*Relevan*)
- 'liu bau qun' = Keywords ini relevan dengan teksnya dikarenakan bisa keyword tersebut merupakan nama korban dari kecelakaan, jadi menurut saya cukup bisa menjadi ringkasan awal terhadap teks. (*Relevan*)

Hasil Code 2

['kasat reskrim polres halmahera tengah iptu taufik saimima mengatakan', 'supervisor komunikasi pt indonesia weda bay industrial park', 'unit mobil truk dt jenis shacman warna merah']

Kesimpulan Code 2

Kesimpulan dari 3 keyword teratas sebagai berikut.

- 'kasat reskrim polres halmahera tengah iptu taufik saimima mengatakan', = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)
- 'supervisor komunikasi pt indonesia weda bay industrial park' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)
- 'unit mobil truk dt jenis shacman warna merah' = Keywords ini relevan dengan teksnya dikarenakan bisa menjadi ringkasan awal terhadap teks serta merupakan poin penting dari teks tersebut. (*Relevan*)

Kesimpulan

- Pada text 1 perbedaannya adalah pada code 1 kandidat yang relevan ada 2, sedangkan pada code 2 semua kandidatnya relevan
- Pada text 2 perbedaannya adalah pada code 1 kandidat yang relevan ada 1, sedangkan pada code 2 semua kandidatnya relevan
- Pada text 3 perbedaannya adalah pada code 1 tidak ada kandidat yang relevan, sedangkan pada code 2 kandidat yang relevan ada 2
- Pada text 4 perbedaannya adalah pada code 1 kandidat yang relevan ada 2, sedangkan pada code 2 semua kandidatnya relevan

Teks	Code 1	Code 2
1	2 (Relevan)	3 (Relevan)
2	1 (Relevan)	3 (Relevan)
3	0 (Relevan)	3 (Relevan)
4	2 (Relevan)	3 (Relevan)

Jadi, berdasarkan hasil output relevansi yang dihasilkan kedua code tersebut adalah parameter pada code 2 lebih optimal dibandingkan pada code 1.

Perbandingan Hasil Beberapa Percobaan Parameter

Parameter 1

```
rake = Rake(  
    min_chars=5,  
    max_words=10,  
    min_freq=2,  
    language_code='id', # 'en'  
    stopwords=stopwordsid2, # {'and', 'of'}  
    lang_detect_threshold=10,  
    max_words_unknown_lang=1,  
    generated_stopwords_percentile=10,  
    generated_stopwords_max_len=1,  
    generated_stopwords_min_freq=1, )
```

Hasil text 1: ['total emisi obligasi' (relevan), 'miliar saham' (tidak relevan), 'terjadi penurunan (tidak relevan)'], (1 Relevan)

Hasil text 2: ['perdagangan' (tidak relevan), 'kenaikan' (tidak relevan), 'terjadi' (tidak relevan)], (0 Relevan)

Hasil text 3: ['orang tua korban' (tidak relevan), 'kolam renang vila' (relevan), 'penjaga vila' (tidak relevan)], (1 Relevan)

Hasil text 4: ['nomor lambung', 'persimpangan jalan', 'jumlah korban'] (0 Relevan)

Parameter 2

```
rake = Rake(  
    min_chars=3,  
    max_words=5,  
    min_freq=1,  
    language_code='id', # 'en'  
    stopwords=stopwordsid2, # {'and', 'of'}  
    lang_detect_threshold=10,  
    max_words_unknown_lang=2,  
    generated_stopwords_percentile=10,  
    generated_stopwords_max_len=2,  
    generated_stopwords_min_freq=1, )
```

Hasil text 1: ['investor asing mencatatkan beli bersih' (relevan), 'aulia noviana utami putri' (relevan), 'indeks harga saham gabungan' (relevan)], (3 Relevan)

Hasil text 2: ['arus deras capital inflow' (relevan), 'berita terkini more menit' (relevan), 'transaksi perdagangan mencapai rp11' (relevan)], (3 Relevan)

Hasil text 3: ['kepala desa cibodas dindin sukaya' (relevan), 'kolam renang vila bellevue' (relevan), 'koran si .' (tidak relevan)], (2 Relevan)

Hasil text 4: ['unit mobil dum truk terjadi' (tidak relevan), 'nomor lambung bermuatan ore nikel' (tidak relevan), 'smelter k perusahaan pt iwip' (tidak relevan)] 0 Relevan

Parameter 3

```
rake = Rake(  
    min_chars=3,  
    max_words=10,  
    min_freq=1,  
    language_code='id', # 'en'  
    stopwords=stopwordsid2, # {'and', 'of'}  
    lang_detect_threshold=10,  
    max_words_unknown_lang=3,  
    generated_stopwords_percentile=10,  
    generated_stopwords_max_len=3,  
    generated_stopwords_min_freq=1, )
```

Hasil text 1: ['investor asing mencatatkan beli bersih' (relevan), 'aulia noviana utami putri' (relevan), 'indeks harga saham gabungan' (relevan)], (3 Relevan)

Hasil text 2: ['ceo indosurya bersinar sekuritas william surya wijaya mengatakan' (relevan), 'arsip foto/mpi/faishal rahman a a a jakarta' (relevan), 'arus deras capital inflow' (relevan)], (3 Relevan)

Hasil text 3: ['kepala desa cibodas dindin sukaya' (relevan), 'kolam renang vila bellevue' (relevan), 'koran si .' (tidak relevan)], (2 Relevan)

Hasil text 4: ['kasat reskrim polres halmahera tengah iptu taufik saimima mengatakan' (relevan), 'supervisor komunikasi pt indonesia weda bay industrial park' (relevan), 'unit mobil truk dt jenis shacman warna merah' (relevan)] (3 Relevan)

Parameter 4

```
rake = Rake(  
    min_chars=5,  
    max_words=5,  
    min_freq=1,  
    language_code='id', # 'en'  
    stopwords=stopwordsid2, # {'and', 'of'}  
    lang_detect_threshold=10,  
    max_words_unknown_lang=2,  
    generated_stopwords_percentile=10,  
    generated_stopwords_max_len=2,  
    generated_stopwords_min_freq=1, )
```

Hasil text 1: ['investor asing mencatatkan beli bersih' (relevan), 'aulia noviana utami putri' (relevan), 'indeks harga saham gabungan' (relevan)], (3 Relevan)

Hasil text 2: ['arus deras capital inflow' (relevan), 'berita terkini more menit' (tidak relevan), 'transaksi perdagangan mencapai rp11' (relevan)], (2 Relevan)

Hasil text 3: ['kepala desa cibodas dindin sukaya' (relevan), 'kolam renang vila bellevue' (relevan), 'koran si .' (tidak relevan)], (2 Relevan)

Hasil text 4: ['unit mobil dum truk terjadi' (tidak relevan), 'smelter k perusahaan pt iwip' (tidak relevan), 'nomor lambung bermuatan ore nikel (tidak relevan)'] (0 Relevan)

Parameter 5

```
rake = Rake(  
    min_chars=5,  
    max_words=10,  
    min_freq=3,  
    language_code='id', # 'en'  
    stopwords=stopwordsid2, # {'and', 'of'}  
    lang_detect_threshold=10,  
    max_words_unknown_lang=3,  
    generated_stopwords_percentile=10,  
    generated_stopwords_max_len=3,  
    generated_stopwords_min_freq=1, )
```

Hasil text 1: ['transaksi', 'triliun', 'tercatat'], (0 Relevan)

Hasil text 2: ['terjadi'], (0 Relevan)

Hasil text 3: [], (0 Relevan)

Hasil text 4: [] (0 Relevan)

Parameter 6

```
rake = Rake(  
    min_chars=5,  
    max_words=10,  
    min_freq=3,  
    language_code='id', # 'en'  
    stopwords=stopwordsid2, # {'and', 'of'}  
    lang_detect_threshold=10,  
    max_words_unknown_lang=3,  
    generated_stopwords_percentile=10,  
    generated_stopwords_max_len=3,  
    generated_stopwords_min_freq=1, )
```

Hasil text 1: ['investor asing mencatatkan beli bersih' (relevan), 'indeks harga saham gabungan' (relevan), 'aulia noviana utami putri' (relevan)], (3 relevan)

Hasil text 2: ['ceo indosurya bersinar sekuritas william surya wijaya mengatakan' (relevan), 'arsip foto/mpi/faisal rahman a a a jakarta' (relevan), 'arus deras capital inflow' (relevan)], (3 relevan)

Hasil text 3: ['kepala desa cibodas dindin sukaya' (relevan), 'kolam renang vila bellevue' (relevan)], (2 relevan)

Hasil text 4: ['kasat reskrim polres halmahera tengah iptu taufik saimima mengatakan' (relevan), 'supervisor komunikasi pt indonesia weda bay industrial park' (relevan), 'unit mobil truk dt jenis shacman warna merah' (relevan)] (3 relevan)

Parameter 7

```
rake = Rake(  
    min_chars=20,  
    max_words=50,  
    min_freq=1,  
    language_code='id', # 'en'  
    stopwords=stopwordsid2, # {'and', 'of'}  
    lang_detect_threshold=10,  
    max_words_unknown_lang=3,  
    generated_stopwords_percentile=10,  
    generated_stopwords_max_len=3,  
    generated_stopwords_min_freq=1, )
```

Hasil text 1: ['investor asing mencatatkan beli bersih' (relevan), 'indeks harga saham gabungan' (relevan), 'aulia noviana utami putri' (relevan)], (3 relevan)

Hasil text 2: ['ceo indosurya bersinar sekuritas william surya wijaya mengatakan' (relevan), 'arsip foto/mpi/faisal rahman a a a jakarta' (relevan), 'arus deras capital inflow' (relevan)], (3 relevan)

Hasil text 3: ['kepala desa cibodas dindin sukaya' (relevan), 'kolam renang vila bellevue' (relevan), 'kabupaten bandung barat' (relevan)], (3 relevan)

Hasil text 4: ['kasat reskrim polres halmahera tengah iptu taufik saimima mengatakan' (relevan), 'supervisor komunikasi pt indonesia weda bay industrial park' (relevan), 'unit mobil truk dt jenis shacman warna merah' (relevan)] (3 relevan)

Kesimpulan Perbandingan Hasil Beberapa Percobaan Parameter

Teks	Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5	Parameter 6	Parameter 7
1	1 (Relevan)	3 (Relevan)	3 (Relevan)	3 (Relevan)	0 (Relevan)	3 (Relevan)	3 (Relevan)
2	0 (Relevan)	3 (Relevan)	3 (Relevan)	3 (Relevan)	0 (Relevan)	3 (Relevan)	3 (Relevan)
3	1 (Relevan)	2 (Relevan)	2 (Relevan)	2 (Relevan)	0 (Relevan)	2 (Relevan)	3 (Relevan)
4	0 (Relevan)	0 (Relevan)	3 (Relevan)	0 (Relevan)	0 (Relevan)	3 (Relevan)	3 (Relevan)
Total	2 Relevan	8 Relevan	11 Relevan	8 Relevan	0 Relevan	11 Relevan	12 Relevan

Kesimpulan dari perbandingan hasil beberapa percobaan parameter yang paling terbaik adalah percobaan parameter ke 7.