

***DOCUMENT RETRIEVAL MENGGUNAKAN MAXIMUM
ENTROPY DAN QUERY EXPANSION***

SKRIPSI

Oleh :
AURELLESIA WARSITO
NIM. 17650092



**JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2020**

***DOCUMENT RETRIEVAL MENGGUNAKAN MAXIMUM
ENTROPY DAN QUERY EXPANSION***

SKRIPSI

**Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)**

**Oleh:
AURELLESIA WARSITO
NIM. 17650092**

**JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2020**

HALAMAN PERSETUJUAN

***DOCUMENT RETRIEVAL MENGGUNAKAN MAXIMUM ENTROPY
DAN QUERY EXPANSION***

SKRIPSI

**Oleh:
AURELLESIA WARSITO
NIM. 17650092**

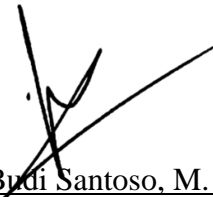
Telah Diperiksa dan Disetujui untuk Diuji
Tanggal : Juni 2021

Dosen Pembimbing I



Dr. Cahyo Crysdiان
NIP. 19740424 200901 1 008

Dosen Pembimbing II



Irwan Budi Santoso, M. Kom
NIP. 19770103 201101 1 004

Mengetahui,
Ketua Jurusan Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Dr. Cahyo Crysdiان
NIP. 19740424 200901 1 008

HALAMAN PENGESAHAN

***DOCUMENT RETRIEVAL MENGGUNAKAN MAXIMUM ENTROPY
DAN QUERY EXPANSION***

SKRIPSI

**Oleh:
AURELLESIA WARSITO
NIM. 17650092**

Telah Dipertahankan di Depan Dewan Penguji Skripsi dan
Dinyatakan Diterima Sebagai Salah Satu Persyaratan untuk
Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal : Juni 2021

Susunan Dewan Penguji

Penguji Utama : Dr. M. Amin Hariyadi, M.T
NIP. 19670118 200501 1 001
Ketua Penguji : A'la Syauqi, M.Kom
NIP. 19771201 200801 1 007
Sekretaris Penguji : Dr. Cahyo Crysdiان
NIP. 19740424 200901 1 008
Anggota Penguji : Irwan Budi Santoso, M. Kom
NIP. 19770103 201101 1 004

Tanda Tangan

()
()
()
()

Mengetahui dan Mengesahkan,
Ketua Jurusan Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Dr. Cahyo Crysdiان
NIP. 19740424 200901 1 008

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini :

Nama : Aurellesia Warsito

NIM : 17650092

Fakultas : Sains dan Teknologi

Jurusan : Teknik Informatika

Judul Skripsi : *Document Retrieval Menggunakan Maximum Entropy dan Query Expansion*

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan data, tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila di kemudian hari terbukti atau dapat dibuktikan Skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 19 Mei 2021
Yang membuat pernyataan,


Aurellesia Warsito
NIM. 17650092

HALAMAN MOTTO

“Allah itu baik.”

HALAMAN PERSEMBAHAN

الْعَلَمِينَ رَبِّ اللَّهِ الْحَمْدُ

Puji syukur kehadiran Allah *subhanahu wa ta'ala*, shalawat serta salam kepada Rasulullah *shalallahu 'alaihi wa sallam*.

Karya ini penulis persembahkan kepada :

Kepada kedua orang tua penulis, Bapak Warsito dan Ibu Siti Khotijah yang senantiasa mendidik, menyayangi, dan mendukung penulis dalam hal apapun, serta menjadi penghibur ketika penulis mengalami masa sulit. Tidak ada kata yang dapat menggambarkan rasa terimakasih penulis kepada mereka, semoga rahmat dan kasih sayang Allah selalu tercurahkan kepada mereka.

Seluruh guru – guru penulis semasa sekolah khususnya para guru Matematika yang telah mengajari dan membimbing penulis dalam memahami materi sampai penulis berada di posisi saat ini.

Seluruh dosen Teknik Informatika, khususnya dosen pembimbing penulis Bapak Dr. Cahyo Crysdiyan dan Bapak Irwan Budi Santoso, M.Kom yang senantiasa membimbing dan membantu memahami materi.

Teman seperjuangan *Unocore* TI 17 UIN Maulana Malik Ibrahim Malang yang selalu memberikan semangat dan do'a kepada penulis, khususnya Aradea Atfal Risdianto yang selalu memotivasi penulis dalam menyelesaikan skripsi ini. Penulis ucapkan terimakasih, semoga Allah membalas kebaikan kalian. آمين

KATA PENGANTAR

السَّلَامُ عَلَيْكُمْ وَرَحْمَةُ اللَّهِ وَبَرَكَاتُهُ

Segala puji bagi Allah *subhanahu wa ta'ala* yang telah memberikan rahmat dan hidayah-Nya serta nikmat lahir dan batin sehingga penulis dapat menyelesaikan skripsi ini. Shalawat serta salam selalu tercurahkan kepada Rasulullah *shalallahu 'alaihi wa sallam* yang telah membimbing umat manusia menuju ke peradaban yang lebih baik dan beradab.

Tujuan dari penyusunan skripsi ini adalah untuk memenuhi salah satu persyaratan untuk bisa menempuh ujian sarjana komputer pada Jurusan Teknik Informatika Fakultas Sains dan Teknologi (SAINTEK) di Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang. Dalam menyelesaikan skripsi ini, banyak pihak yang telah memberikan bantuan secara moril, nasihat, semangat maupun materiil. Penulis sampaikan terimakasih yang sedalam-dalamnya kepada :

1. Prof. Dr. Abdul Haris, M.Ag selaku rektor Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.
2. Dr. Sri Hariani, M.Si selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.
3. Dr. Cahyo Crysdiyan selaku Ketua Jurusan Teknik Informatika sekaligus Dosen Pembimbing I yang telah membimbing penyusunan skripsi ini hingga selesai.

4. Irwan Budi Santoso, M.Kom selaku Dosen Pembimbing II yang telah membimbing penyusunan skripsi ini hingga selesai.
5. Dr. Amin Hariyadi, M.T dan A'la Syauqi, M.Kom selaku Dosen Penguji yang telah memberikan banyak saran perbaikan kepada penulis.
6. Citra Fidya Atmalia, S.H selaku admin Jurusan Teknik Informatika yang senantiasa memberikan informasi serta arahan dalam penyusunan skripsi.
7. Orang tua tercinta yang telah memberikan dukungan serta do'a sehingga penulis dapat menyelesaikan skripsi.
8. Teman seperjuangan yang selalu memberikan motivasi dan dukungan kepada penulis sehingga penulis dapat menyelesaikan skripsi dengan baik dan dapat lulus tepat waktu.
9. Teman-teman Unocore yang selalu memberikan dukungan dan bantuan kepada penulis selama menjadi mahasiswa.
10. Semua pihak yang terlibat dalam perjuangan penulis untuk menyelesaikan skripsi.

Penulis menyadari bahwa skripsi ini masih memiliki banyak kekurangan dan kesalahan, untuk itu penulis menerima segala kritik dan saran serta semoga skripsi ini senantiasa memberikan manfaat kepada para pembaca.

Malang, 18 Mei 2021

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN KEASLIAN TULISAN	iv
HALAMAN MOTTO	iv
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
ABSTRAK	xiii
ABTSRACT.....	xiv
الملخص.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Pernyataan Masalah	6
1.3 Tujuan Penelitian	6
1.4 Batasan Masalah	6
1.5 Manfaat Penelitian	7
BAB II STUDI PUSTAKA.....	8
2.1 <i>Natural Language Processing</i>	8
2.2 <i>Query Expansion</i>	9
2.3 <i>Term Probability</i>	10
2.4 <i>Shannon Entropy</i>	11
2.5 <i>Information Retrieval</i>	12

2.6 <i>Theoretical Framework</i>	13
BAB III METODOLOGI PENELITIAN.....	18
3.1 Pengumpulan Data.....	18
3.2 Desain Sistem	18
3.2.1 <i>Preprocessing</i>	19
3.2.2 <i>Query Expansion</i>	23
3.2.3 <i>Indexing</i>	26
3.2.4 <i>Term Probability</i>	28
3.2.5 <i>Shannon's Entropy</i>	31
BAB IV UJI COBA DAN PEMBAHASAN	34
4.1 Langkah Uji Coba.....	34
4.1.1 <i>Preprocessing</i>	34
4.1.2 <i>Query Expansion</i>	35
4.1.3 <i>Indexing</i>	36
4.1.4 <i>Term Probability</i>	37
4.1.5 <i>Shannon's Entropy</i>	39
4.2 Hasil Uji Coba	40
4.3 Pembahasan	44
BAB V KESIMPULAN DAN SARAN.....	49
5.1 Kesimpulan	49
5.2 Saran	50
DAFTAR PUSTAKA	51
LAMPIRAN.....	54

DAFTAR GAMBAR

Gambar 2.1 <i>Theoretical Framework</i>	14
Gambar 3.1. <i>Desain Sistem</i>	19
Gambar 3.2. <i>Tahap Preprocessing</i>	20
Gambar 3.3. <i>Flowchart Stemming</i>	22
Gambar 3.4. <i>Flowchart Query Expansion</i>	23
Gambar 3.6. <i>Flowchart Indexing</i>	27
Gambar 3.7. <i>Source Code Indexing</i>	28
Gambar 3.8. <i>Flowchart Term Probability</i>	29
Gambar 3.9 <i>Source Code Term Probability</i>	30
Gambar 3.10 <i>Flowchart Entropy</i>	32
Gambar 3.11. <i>Source Code Entropy</i>	33
Gambar 3.12. <i>Source Code Ranking</i>	33

DAFTAR TABEL

Tabel 3.1 Contoh Hasil <i>Case Folding</i>	20
Tabel 3.2. Contoh Hasil <i>Filtering</i>	21
Tabel 3.3. Contoh Hasil <i>Tokenizing</i>	21
Tabel 3.4. Contoh Hasil <i>Stemming</i>	22
Tabel 3.5. Contoh Kamus Sinonim.....	24
Tabel 3.6. Contoh <i>Query Expansion</i>	24
Tabel 4.1. Contoh Hasil <i>Preprocessing Query</i>	34
Tabel 4.2. Contoh Hasil <i>Preprocessing</i> Dokumen	35
Tabel 4.3. Contoh Kamus Sinonim.....	35
Tabel 4.4. Contoh Hasil <i>Query Expansion</i>	36
Tabel 4.5. Contoh Hasil <i>Indexing</i>	37
Tabel 4.6. Contoh Hasil Perhitungan <i>Term Probability</i>	38
Tabel 4.8. Daftar <i>Query</i> Uji Coba.....	40
Tabel 4.9. Hasil Uji Coba Tanpa Menggunakan <i>Query Expansion</i>	41
Tabel 4.10. Hasil Uji Coba Menggunakan <i>Query Expansion</i>	42

ABSTRAK

Warsito, Aurellesia. 2021. *Document Retrieval Menggunakan Maximum Entropy dan Query Expansion*. Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing : (I) Dr. Cahyo Crysdiyan (II) Irwan Budi Santoso, M.Kom.

Kata Kunci : *User's query, Query Expansion, Shannon's Entropy, Term Probability, Information Retrieval.*

Indonesia merupakan negara hukum yang menjalankan kehidupan berbangsa, bernegara serta bermasyarakat berdasarkan peraturan perundang-undangan. Hal ini membuat masyarakat harus memiliki pengetahuan mengenai peraturan di Indonesia agar tidak melakukan pelanggaran hukum yang dapat merugikan banyak pihak. Maka dari itu diperlukan sebuah sarana berupa aplikasi untuk mencari dokumen-dokumen peraturan perundang-undangan yang dapat digunakan oleh banyak pihak menggunakan kata kunci atau *query*. Dalam penelitian ini aplikasi pencarian dokumen perundang-undangan dibangun menggunakan metode *Shannon's entropy* dan *query expansion*. *Shannon's entropy* digunakan untuk menghitung tingkat kecocokan antara *query* dari *user* dengan isi dari setiap dokumen perundang-undangan. Semakin tinggi nilai *entropy* dari sebuah dokumen maka semakin besar pula tingkat korelevanan dokumen dengan *query*. Metode ini digunakan karena dapat mengukur relevansi isi dokumen PDF dengan *query* dari *user* berdasarkan probabilitas *term* yang terdapat dalam dokumen. Tujuan dari penelitian ini adalah untuk mengukur perbandingan kinerja metode tersebut tanpa atau dengan menggunakan *query expansion* dengan cara melakukan uji evaluasi sistem dari aspek *recall*, *precision*, *f-measure* dan *accuracy*. Hasil penelitian menunjukkan bahwa metode *Shannon's entropy* tanpa menggunakan *query expansion* memiliki nilai *precision* sebesar 41,33%, nilai *recall* sebesar 41,33%, nilai *f-measure* sebesar 41,33% dan nilai *accuracy* sebesar 99,41% dalam mengembalikan dokumen sejumlah 10 teratas terhadap masing-masing *query*. Sedangkan metode *Shannon's entropy* dengan menggunakan *query expansion* menghasilkan nilai *precision* sebesar 43,67%, nilai *recall* sebesar 43,67%, nilai *f-measure* sebesar 43,67% dan nilai *accuracy* sebesar 99,44%. Nilai *precision*, *recall*, dan *f-measure* meningkat sebesar 2,34% dan nilai *accuracy* meningkat sebesar 0,3%. Berdasarkan hasil pengujian tersebut, dapat disimpulkan bahwa metode *Shannon's entropy* menggunakan *query expansion* memiliki kinerja lebih baik dibanding tanpa menggunakan *query expansion*.

ABSTRACT

Warsito, Aurellesia. 2021. *Document Retrieval Using Maximum Entropy and Query Expansion*. Undergraduate Thesis. Department of Computer Science, Faculty of Science and Technology, Islamic State University of Maulana Malik Ibrahim of Malang. Supervisor : (I) Dr. Cahyo Crys dian (II) Irwan Budi Santoso, M.Kom.

Keywords : *User's query, Query Expansion, Shannon's Entropy, Term Probability, Information Retrieval.*

Indonesia is a constitutional state that carries out the life of the nation, state and society based on statutory regulations. This makes the public must have knowledge of regulations in Indonesia so as not to commit violations of the law that can harm many parties. Therefore we need a facility in the form of an application to search for statutory documents that can be used by many parties using keywords or queries. In this study, the statutory document search application was built using Shannon's entropy method and query expansion. Shannon's entropy is used to calculate the match level between a user's query and the content of each statutory document. The higher the entropy value of a document, the greater the relevance of the document to the query. This method is used because it can measure the relevance of the contents of a PDF document with a query from the user based on the probability terms contained in the document. The purpose of this study is to measure the performance comparison of these methods without or by using query expansion by performing a system evaluation test from the aspects of recall, precision, f-measure and accuracy. The results showed that the Shannon's entropy method without using query expansion has a precision value of 41.33%, a recall value of 41.33%, an f-measure value of 41.33% and an accuracy value of 99.41% in returning documents of 10. top of each query. Meanwhile, Shannon's entropy method using query expansion produces a precision value of 43.67%, a recall value of 43.67%, an f-measure value of 43.67% and an accuracy value of 99.44%. The precision, recall, and f-measure values increase by 2.34% and the accuracy value increases by 0.3%. Based on the results of these tests, it can be concluded that Shannon's entropy method using query expansion has better performance than without using query expansion.

الملخص

وارسيتو ، أوريليسيا. ٢٠٢١. استرداد المستند باستخدام أقصى إنتروبيا وتوسيع الاستعلام . أطروحة. قسم المعلوماتية ، كلية العلوم والتكنولوجيا ، جامعة الدولة الإسلامية مولانا مالك إبراهيم مالانج. المستشار: (١) د. كاهيو كريسيديان (٢) إيروان بودي سانتوسو ، إم كوم.

الكلمات الرئيسية: استعمال المستخدم ، توسيع الاستعلام ، إنتروبيا شانون ، احتمالية المصطلح ، استرجاع المعلومات.

إندونيسيا دولة قانونية تمارس حياة الأمة والدولة والمجتمع على أساس القوانين واللوائح. هذا يجعل الجمهور على دراية باللوائح في إندونيسيا حتى لا ينتهك القانون الذي يمكن أن يضر العديد من الأطراف. لذلك نحن بحاجة إلى أداة في شكل تطبيق للبحث عن وثائق التشريع التي يمكن استخدامها من قبل العديد من الأطراف باستخدام الكلمات الرئيسية أو الاستعلامات. في هذه الدراسة ، تم إنشاء تطبيق البحث عن المستندات القانونية باستخدام طريقة شانون الانتروبيا وتوسيع الاستعلام. تُستخدم إنتروبيا شانون لحساب درجة التطابق بين الاستعلام من المستخدم ومحتويات كل مستند قانوني. كلما زادت قيمة الانتروبيا للمستند ، زاد مستوى ملاءمة المستند للاستعلام. تُستخدم هذه الطريقة لأنها يمكن أن تقيس مدى ملاءمة محتويات مستند PDF باستخدام من المستخدم بناءً على شروط الاحتمال الواردة في المستند. الغرض من هذه الدراسة هو قياس مقارنة أداء الطريقة بدون أو باستخدام توسيع الاستعلام عن طريق إجراء اختبار تقييم النظام من جوانب الاسترجاع والدقة والقياس f والدقة. أظهرت النتائج أن طريقة إنتروبيا شانون بدون استخدام توسيع الاستعلام لها قيمة دقة تبلغ 41.33% وقيمة استدعاء 41.33% وقيمة قياس f 41.33% وقيمة دقة 99.41% في إرجاع 10 مستندات. أعلى لكل منها استفسار. بينما تنتج طريقة إنتروبيا شانون باستخدام توسيع الاستعلام قيمة دقة تبلغ 43.67% وقيمة استدعاء 43.67% وقيمة قياس f 43.67% وقيمة دقة 99.44%. زادت قيمة الدقة والاسترجاع والقياس f بنسبة 2.34% وزادت قيمة الدقة بنسبة 0.3%. استنادًا إلى نتائج الاختبار ، يمكن استنتاج أن طريقة الانتروبيا الخاصة بشانون باستخدام توسيع الاستعلام لها أداء أفضل من دون استخدام توسيع الاستعلام.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara hukum, seperti itulah bunyi Pasal 1 Ayat 3 dalam UUD 1945 Republik Indonesia. Hal ini menunjukkan bahwa semua aspek kehidupan di Indonesia baik kehidupan bermasyarakat atau kehidupan bernegara harus dilandasi dengan hukum. Menurut Kamus Besar Bahasa Indonesia, hukum adalah peraturan atau adat yang secara resmi mengikat, yang dikukuhkan oleh penguasa atau pemerintah. Immanuel Kant mengajarkan tentang teori negara hukum di mana tujuan negara adalah menegakkan hak warganya. Peran penegakan supremasi hukum juga terdapat dalam upaya pemerintah untuk mencapai keteraturan dan ketertiban yang dibutuhkan.

Nugraheny (2019) menjelaskan berdasarkan Data Kependudukan Semester I 2020 yang dirilis oleh Kementerian Dalam Negeri melalui Direktorat Jenderal Kependudukan dan Pencatatan Sipil (Dukcapil), jumlah penduduk Indonesia per 30 Juni mencapai 268.583.016 jiwa. Jumlah penduduk Indonesia yang besar berpotensi meningkatkan pelanggaran hukum. Pelanggaran hukum yang dilakukan oleh masyarakat dapat terjadi di berbagai sektor dengan berbagai jenis tindak pidana. Menurut Badan Pusat Statistik (2019), data yang dihimpun oleh Subdirektorat Statistik Politik dan Keamanan menunjukkan terdapat 294.281 pelanggaran yang terjadi di tahun 2019.

Hubungan antara hukum dan masyarakat terbilang masih minim. Hal ini dapat dilihat dari banyaknya tindak pidana yang dilakukan oleh masyarakat. Mukaram (2019) memaparkan warga Kalimantan Barat harus berurusan dengan Polda Kalimantan Barat karena melakukan penghinaan terhadap lambang negara dengan mengunggah dan mengganti Pancasila menjadi Pancasila. Tindakan seperti ini jelas melanggar peraturan perundang-undangan di Indonesia dan dijerat pasal 45A ayat (1) dan pasal 28 ayat (1) UU No 19 tahun 2016 atas perubahan UU No 11 tahun 2008 Tentang Informasi dan Transaksi Elektronik dan pasal 68 UU No 24 Tahun 2009 tentang Bendera, Bahasa dan Lambang Negara serta Lagu Kebangsaan. Kasus lain yang melanggar peraturan perundangan adalah seperti yang dilansir oleh CNN Indonesia (2020), yaitu perbuatan salah satu Youtuber asal Bandung yang membuat konten prank memberikan sembako berisi sampah dan batu kepada waria di Bandung sehingga dijerat dengan Pasal 43 ayat 3 UU ITE tentang penghinaan atau pencemaran nama baik melalui media elektronik serta Pasal 36 dan Pasal 51 ayat 2 UU ITE Nomor 11 Tahun 2008.

Tindak pidana seperti yang disebutkan di atas memiliki banyak faktor penyebab, salah satunya adalah kurangnya pengetahuan masyarakat tentang perundang-undangan di Indonesia sehingga masyarakat merasa bebas dari hukum ketika melakukan tindakan yang tidak benar di mata negara. Ketidaktahuan masyarakat tentang hak dan kewajiban sebagai warga negara membuat pelanggaran hukum sering terjadi dan merugikan masyarakat itu sendiri. Sementara itu, pemerintah telah menciptakan dan memberlakukan berbagai macam peraturan perundang-undangan guna mengatur dan memberikan dasar

untuk setiap perilaku masyarakat agar tetap pada jalur hukum yang berlaku. Menurut Walukow (2013), peraturan yang ditetapkan oleh pemerintah berlaku untuk semua kalangan masyarakat dan semua orang dianggap sama di depan hukum, sesuai dengan teori *Equality Before The Law*.

Allah *subhanahu wa ta'ala* berfirman dalam surah An-Nisa ayat 59 :

يَا أَيُّهَا الَّذِينَ آمَنُوا أَطِيعُوا اللَّهَ وَأَطِيعُوا الرَّسُولَ وَأُولَى الْأَمْرِ مِنْكُمْ فَإِنْ تَنَازَعْتُمْ فِي شَيْءٍ فَرُدُّوهُ إِلَى اللَّهِ وَالرَّسُولِ إِنْ كُنْتُمْ تُؤْمِنُونَ بِاللَّهِ وَالْيَوْمِ الْآخِرِ ذَلِكَ خَيْرٌ وَأَحْسَنُ تَأْوِيلًا

“Hai orang-orang yang beriman, taatilah Allah dan taatilah Rasul (Nya), dan ulil amri di antara kamu. Kemudian jika kamu berlainan pendapat tentang sesuatu, maka kembalikanlah ia kepada Allah (Al-Qur'an) dan Rasul (sunnahnya), jika kamu benar-benar beriman kepada Allah dan hari kemudian. Yang demikian itu lebih utama (bagimu) dan lebih baik akibatnya”.

Ayat di atas menjelaskan bahwa Allah *subhanahu wa ta'ala* memerintahkan untuk taat dan patuh kepada ulil amri, yaitu pemimpin di mana dalam hal ini adalah pemerintah. Taat yang diperintahkan oleh Allah *subhanahu wa ta'ala* merupakan kewajiban selama pemerintah tidak memerintah kepada keburukan. Peraturan perundang-undangan yang dibuat dan ditetapkan oleh pemerintah tentunya bukan untuk tujuan keburukan melainkan untuk mengatur dan mengarahkan masyarakat sehingga dapat tercapai kehidupan yang aman dan damai.

Lemahnya kesadaran hukum masyarakat juga menjadi faktor terjadinya pelanggaran. Di Indonesia sendiri masyarakat masih sering melakukan pelanggaran tanpa ada rasa takut akan sanksi hukum yang diberikan oleh pihak berwenang. Membuang sampah di sungai adalah contoh pelanggaran yang justru menjadi kebiasaan di tengah masyarakat dan dianggap wajar. Peraturan dan hukum yang dibuat oleh pemerintah sudah selayaknya berlaku untuk semua aspek kehidupan sehari-hari termasuk membuang sampah. Perilaku masyarakat yang mencerminkan kesadaran hukum yang rendah dapat merugikan banyak pihak dan menjadi penghambat tujuan negara itu sendiri.

Menurut Balitbangkumham (2016), nilai-nilai yang ada pada diri manusia tentang hukum yang ada merupakan makna dari kesadaran hukum. Masyarakat Indonesia sudah seharusnya memiliki kesadaran hukum yang baik dan tunduk kepada peraturan yang telah dibuat pemerintah tanpa adanya paksaan dari pihak manapun. Semakin tinggi tingkat sadar hukum suatu masyarakat maka akan semakin baik juga tatanan sebuah negara karena semua segi kehidupan berjalan dengan baik dan aman. Pendidikan tentang pentingnya hukum juga perlu diterapkan kepada seluruh lapisan masyarakat untuk meningkatkan kesadaran hukum, khususnya siswa dan mahasiswa yang berperan sebagai penerus bangsa dan tentunya memiliki akses lebih besar untuk belajar mengenai hukum ketika di sekolah atau di perguruan tinggi.

Di sisi lain, jenjang pendidikan bukanlah menjadi faktor utama seseorang untuk mengerti hukum. Banyak pelanggar hukum yang justru memiliki riwayat

pendidikan tinggi. Ini membuktikan bahwa pendidikan dan latar belakang seseorang tidak menentukan tingkat sadar hukum yang dimiliki. Maka semua lapisan masyarakat dari berbagai latar belakang memiliki kesempatan yang sama untuk belajar dan mengerti tentang hukum yang berlaku di Indonesia.

Pemerintah telah memberikan akses yang lebih mudah kepada masyarakat untuk mengetahui peraturan perundang-undangan melalui situs yang disediakan oleh masing-masing kementerian yang bernama JDIH (Jaringan Dokumentasi dan Informasi Hukum). Menurut Peraturan Menteri Kominfo No. 20 Tahun 2013 tentang Jaringan Dokumentasi dan Informasi Hukum, fungsi dari JDIH adalah sebagai wadah pendayagunaan bersama dokumentasi hukum dan saran pemberian informasi hukum secara lengkap, mudah, akurat, dan tepat. Setiap kementerian menyediakan situs JDIH masing-masing, sehingga pencarian dokumen hukum harus mengakses situs kementerian yang dituju terlebih dahulu. Hal ini tentunya tidak cukup efisien mengingat masyarakat awam yang ingin mencari dokumen hukum tidak tahu menahu tentang jenis-jenis hukum dan bidang kementerian.

Masyarakat khususnya pelajar dan mahasiswa membutuhkan sebuah sarana yang mudah untuk belajar tentang hukum. Maka perlu dibangun sebuah sistem yang dapat memfasilitasi masyarakat dalam mencari dan menelusuri dokumen peraturan perundang-undangan. Kemajuan teknologi membuat setiap orang mampu untuk mencari informasi dengan mudah. Sistem yang perlu dibangun adalah sistem di mana setiap orang dapat mencari dokumen peraturan perundang-undangan sesuai dengan topik-topik tertentu tanpa harus memikirkan

kementerian yang memproduksi peraturan tersebut. Dengan begitu diharapkan masyarakat dapat mengakses dan mempelajari peraturan perundang-undangan dengan mudah sehingga dapat meminimalisir kasus kriminalitas di Indonesia dan menciptakan kehidupan bernegara yang aman dan damai.

1.2 Pernyataan Masalah

1. Seberapa tinggi *accuracy*, *precision*, *recall*, dan *f-measure* metode *Shannn's entropy* dalam melakukan *document retrieval* berdasarkan *query user*?
2. Seberapa tinggi peningkatan *performance Shannon's entropy* jika dilakukan proses *query expansion* ditinjau dari parameter *accuracy*, *precision*, *recall*, dan *f-measure* ?

1.3 Tujuan Penelitian

1. Mengukur *accuracy*, *precision*, *recall*, dan *f-measure* metode *Shannon's entropy* dalam melakukan *document retrieval* berdasarkan *query user*.
2. Mengukur peningkatan *performance Shannon's entropy* jika dilakukan proses *query expansion* ditinjau dari parameter *accuracy*, *precision*, *recall*, dan *f-measure*.

1.4 Batasan Masalah

1. Data yang digunakan adalah dokumen perundang-undangan yang berjumlah dua ribu dari tiga belas bidang kementerian yaitu bidang Agama, Agraria, Kelautan, Pendidikan dan Budaya, Kesehatan,

Ketenagakerjaan, Keuangan, Komunikasi dan Informasi, Lingkungan Hidup, Pertahanan, Perhubungan, Politik dan Hukum, serta Riset Penelitian dan Perguruan Tinggi.

2. Dokumen perundang-undangan adalah dokumen berbahasa Indonesia dengan format pdf.
3. Sistem yang dibangun merupakan sistem berbasis web.

1.5 Manfaat Penelitian

1. Mempermudah mahasiswa jurusan hukum untuk melakukan pencarian dokumen perundang-undangan di Indonesia.
2. Mempermudah pelajar Sekolah Menengah Pertama dan Sekolah Menengah Atas dalam melakukan pencarian dokumen perundang-undangan untuk memenuhi tugas sekolah.

BAB II

STUDI PUSTAKA

2.1 *Natural Language Processing*

Natural Language Processing (NLP) tidak hanya digunakan dalam dunia informatika, Becker *et al.* (2019) melakukan sebuah penelitian yang menerapkan konsep NLP pada bidang kesehatan. NLP yang digunakan dalam penelitian ini bertujuan untuk mengidentifikasi informasi penyakit dari pasien kanker kolon sesuai dengan pedoman standar dan mencatat hasil identifikasi menggunakan bahasa medis terpadu untuk selanjutnya digunakan oleh dokter sebagai bahan evaluasi. Algoritma yang dikembangkan adalah algoritma yang dapat memetakan entitas dengan konsep UMLS untuk Bahasa Jerman. Selama proses *training*, sebuah basis data khusus menampung sinonim *terminology* kesehatan Bahasa Jerman juga dibangun. Data yang digunakan pada penelitian ini berupa 2.513 catatan medis Jerman dan menghasilkan presisi rata-rata sebesar 82,05%, *recall* sebesar 82,45% dan *f-measure* masing-masing 81,81%.

Otomatisasi merupakan sebuah cara untuk menggantikan tenaga manusia ke tenaga mesin guna mempermudah pekerjaan manusia. Terdapat berbagai metode dalam menciptakan sebuah otomatisasi, salah satunya adalah NLP. Ly *et al.* (2018) dalam artikel penelitiannya yang berjudul *Evaluation of Natural Language Processing (NLP) Systems to Annotate Drug Product Labeling with MedDRA Terminology* menjelaskan bagaimana NLP dapat bekerja menciptakan otomatisasi. Dalam penelitian ini, uji coba dilakukan pada tiga teknik NLP, yaitu

ETHER, I2E, dan MetaMap untuk dapat mengekstrak istilah *adverse events* (AE) atau efek samping pada sebuah label obat yang kemudian diterjemahkan menggunakan MedDRA secara otomatis. MedDRA sendiri adalah sebuah standar untuk penulisan AE dalam dunia farmasi. Dalam penelitian ini dikembangkan sebuah pedoman anotasi berbasis farmakovigilans untuk mengekstrak istilah AE. Pemetaan manual dilakukan oleh ahli farmakovigilans sebagai perbandingan *output* sistem. Hasil dari penelitian ini adalah NLP mampu melakukan otomatisasi pemetaan yang baik.

Penelitian yang dilakukan oleh Ramaswamy dan DeClerk (2018) menjelaskan tentang penggunaan NLP untuk menganalisa minat pelanggan di pasaran melalui media sosial berbasis teks. Metode pada penelitian ini menggunakan *Natural Language Understanding* (NLU) untuk melakukan *semantic tagging* atau penandaan kata-kata yang memiliki makna yang berhubungan dengan produksi mobil sebagai produk perusahaan. Pembuatan catatan acak (*custom annotation*) untuk setiap kata yang berhubungan secara semantik dari *review* mobil. Serta menciptakan ontologi bisnis untuk menentukan istilah yang berhubungan dengan produksi mobil. Hasil dari penelitian ini adalah NLP berhasil menjelaskan minat pelanggan di pasaran mengenai mobil yang akan menjadi produk utama sebuah perusahaan.

2.2 Query Expansion

Singh dan Sharan (2015) melakukan penelitian yang menggunakan *pseudo-relevance feedback* (PRF) yang merupakan salah satu teknik dalam *query*

expansion yang dapat meningkatkan kinerja dari *information retrieval*. Hasil dari penelitian ini adalah pendekatan baru mengenai PRF yang dapat meningkatkan kinerja sistem *information retrieval* secara signifikan di atas metode *individual terms selection* dan metode *relate state-of-the-art*.

Dalam penelitian yang dilakukan oleh Fikri (2019), metode *query expansion* digunakan bersama dengan metode *dice similarity* dan *cosine similarity* untuk pencarian *ayatul ahkam* dalam terjemah Al-Qur'an Bahasa Indonesia. *Query expansion* dilakukan melalui keterkaitan kata pada *query* dengan gugusan kata dari thesaurus Bahasa Indonesia. Penelitian ini membandingkan kedua metode tanpa *query expansion* dan dengan *query expansion*. Hasil dari penelitian ini adalah nilai *recall*, *precision*, dan *f-measure* metode dengan *query expansion* lebih tinggi dari metode tanpa *query expansion*.

2.3 Term Probability

Penentuan waktu di mana suatu peristiwa dapat berlangsung di masa depan merupakan masalah yang dipelajari dengan baik di sejumlah disiplin ilmu dan teknik. Lebih dari lima puluh tahun, peneliti mencoba untuk menetapkan metode yang memadai untuk mengkarakterisasi perilaku sistem dinamis pada waktunya. Penelitian yang dilakukan oleh Acuna-ureta *et al.* (2020) yang berjudul *Computation of Time Probability Distribution for The Occurance of Uncertain Future Events* memberikan formalisasi matematis yang ketat untuk masalah komputasi probabilitas terjadinya peristiwa masa depan yang tidak pasti baik dalam proses stokastik waktu diskrit dan kontinu. Penelitian ini focus pada

aplikasi waktu diskrit dengan menunjukkan bagaimana menghitung ukuran probabilitas tersebut dan memvalidasinya dengan membandingkan hasil yang diperoleh dengan simulasi Monte Carlo.

Penelitian yang berjudul *Estimating Product-Choice Probabilities from Recency and Frequency of Page Views* yang dilakukan oleh Iwagana *et al.* (2016) menginvestigasi hubungan antara tampilan halaman pelanggan dan probabilitas pilihan produk mereka di situs *e-commerce*. Untuk tujuan itu, dibuat tabel probabilitas yang berisi probabilitas pilihan produk terkini dan kombinasi frekuensi dari setiap tampilan halaman pelanggan sebelumnya. Untuk mengurangi kesalahan estimasi saat sampel pelatihan sedikit, dikembangkan model optimasi untuk memperkirakan probabilitas pilihan produk yang memenuhi batas monotonisitas, konveksitas, dan konkavitas sehubungan dengan keterkinian dan frekuensi. Hasil komputasi menunjukkan bahwa metode tersebut memiliki keunggulan yang jelas dibandingkan regresi logistic dan *support vector machine* berbasis kernel.

2.4 Shannon Entropy

Banyak algoritma *clustering* telah dikembangkan, sementara kebanyakan dari mereka tidak dapat memproses objek dalam ruang atribut numerik dengan nilai yang hilang. Kebanyakan dari algoritma-algoritma tersebut, jumlah *cluster* harus ditentukan secara manual dan hasil *clustering* sensitif terhadap urutan input dari objek yang akan dikelompokkan. Hal ini dapat membatasi penerapan *clustering* dan mengurangi kualitas *clustering*. Untuk mengatasi masalah ini,

penelitian yang dilakukan oleh Chen dan Wang (2006) dalam artikel penelitian berjudul *Rough Set-Based Clustering with Refinement Using Shannon's Entropy Theory* menyajikan algoritma *clustering* yang ditingkatkan berdasarkan *rough set* (RS) atau himpunan kasar dan teori entropi. Ini bertujuan untuk menghindari penentuan jumlah *cluster* secara manual, dan pengenalan *clustering* dalam ruang atribut numerik dan nominal yang mirip untuk menggantikan indeks jarak. Pada saat yang sama, teori *rough set* menyediakan algoritma dengan fungsi untuk menangani ketidakpastian dalam analisis data. Entropi Shannon digunakan untuk menyempurnakan hasil *clustering* dengan menetapkan bobot relative ke himpunan atribut sesuai dengan nilai entropi yang sama. Hasil penelitian menunjukkan bahwa efisiensi dan kualitas algoritma *clustering* ini meningkat.

Distribusi tegangan geser pada dasar sungai alluvial dan tepian merupakan salah satu parameter penting dalam analisis stabilitas saluran. Penelitian yang dilakukan oleh Kazemian-Kale-Kale *et al.* (2019) berjudul *The Uncertainty of the Shannon Entropy Model for Shear Stress Distribution in Circular Channels* menyajikan metode analisis ketidakpastian yang didasarkan pada kerangka *Bayesian Forecasting System* (BFS) untuk mengevaluasi model entropi Shannon untuk prediksi distribusi tegangan geser.

2.5 Information Retrieval

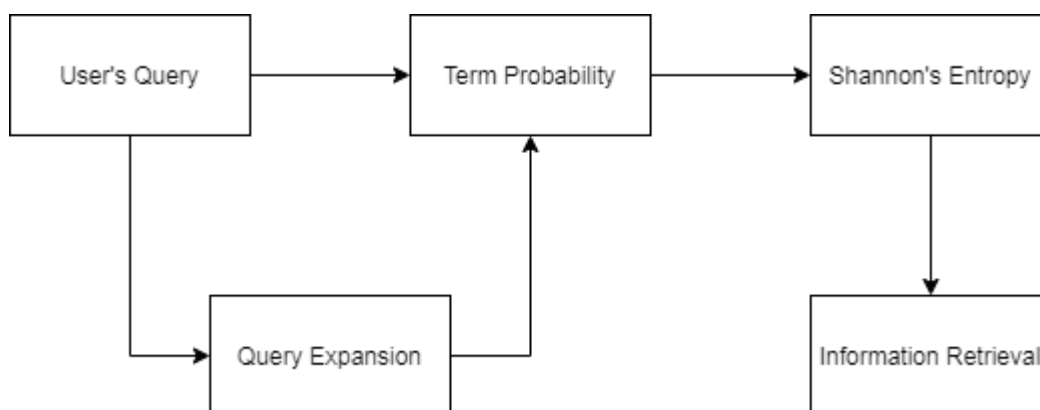
Pohon wavelet adalah sebuah struktur data yang sering digunakan untuk mempresentasikan urutan, permutasi, dan proses lain yang membutuhkan struktur data yang ringkas. Pada artikel penelitian yang dilakukan oleh Gagie *et al.* (2012)

menjelaskan tentang penggunaan pohon wavelet untuk berbagai bentuk kueri yang selanjutnya diimplementasikan pada sistem temu kembali informasi (*information retrieval*). Hasilnya adalah pohon wavelet dan kueri-kueri yang digunakan dapat mengembalikan informasi dengan baik tanpa menambah kebutuhan kapasitas.

Selanjutnya penelitian yang dilakukan oleh Jatnika *et al.* (2019) yang menghitung nilai kemiripan antar kata-kata dalam Bahasa Inggris menggunakan teknik representasi kata. Word2vec adalah model yang digunakan untuk merepresentasikan kata ke bentuk vektor. Studi ini menggunakan 320.000 artikel Wikipedia Bahasa Inggris dan Cosine Similarity untuk mencari nilai kemiripan. Hasil dari penelitian ini adalah semakin banyak teks yang diukur dan semakin tinggi dimensi vektor maka nilai kemiripan akan semakin tinggi pula, namun tidak menutup kemungkinan jika semakin banyak teks yang dibandingkan maka semakin rendah pula nilai kemiripannya karena jumlah yang terlalu besar.

2.6 Theoretical Framework

Menurut Kivunja (2018), *theoretical framework* terdiri dari teori-teori yang diungkapkan oleh para ahli di bidang yang diteliti, yang digunakan untuk analisis data dan interpretasi hasil penelitian. Ini adalah cetak bitu yang sering dipakai oleh peneliti untuk membangun rumah sendiri atau pertanyaan penelitian. *Theoretical framework* berfungsi sebagai fondasi di mana penelitian dibangun. Berdasarkan pengertian tersebut, maka *theoretical framework* yang dihasilkan dari studi literatur ditampilkan pada Gambar 2.1.



Gambar 2.1 *Theoretical Framework*

Pemahaman *query* adalah poin utama dalam pengembalian informasi, yang bertujuan untuk memprediksi hasil pencarian sesuai dengan *query* pencarian yang diberikan oleh pengguna. Menurut Zang *et al.* (2020), keuntungan dalam memahami *user's query* adalah mencapai kinerja yang lebih baik dalam perumusan ulang *query*, pengurangan *query*, prediksi *query* dan saran *query*, serta meningkatkan pemodelan relevansi dokumen berdasarkan maksud dari *query* pencarian. *User's query* merupakan pernyataan berupa kata, frasa, atau kalimat yang memiliki makna yang dimasukkan oleh *user* ke dalam sistem untuk mendapatkan informasi atau dokumen yang sesuai dengan makna *query* tersebut.

Setelah *user* memasukkan *query*, selanjutnya terdapat *text preprocessing*. *Text preprocessing* adalah mengubah file teks mentah, menjadi urutan unit yang dapat didefinisikan secara linguistik dengan baik. Menurut Behzadi (2015), *text preprocessing* adalah bagian utama dari setiap sistem pemrosesan teks, karena karakter, kata, dan kalimat yang diidentifikasi pada tahap ini adalah unit dasar yang diteruskan ke semua tahapan pemrosesan lainnya.

Selanjutnya, *query expansion* adalah sebuah metode untuk meningkatkan jumlah dokumen yang relevan terhadap suatu *query* dengan cara menambahkan kata lain pada sebuah *query* yang memiliki makna yang sama. Adanya *query expansion* membuat dokumen yang ditemukan di dalam repositori menjadi lebih banyak. Menurut Yousef *et al.* (2010) metode *query expansion* terdiri dari *manual query expansion (MQE)* dan *automatic query expansion (AQE)*. *Manual query expansion* memungkinkan *user* untuk melakukan penambahan kata secara manual melalui input. Sedangkan, *automatic query expansion* memiliki beberapa teknik di antaranya adalah *global analysis (GA)* yang membangun *query* menggunakan istilah-istilah yang bermakna sama dari kamus, *local analysis (LA)* memperluas *query* menggunakan informasi pada dokumen yang dikembalikan menggunakan *query* awal, dan *local context analysis (LCA)* yang menggabungkan kelebihan dari teknik *global analysis* dan *local analysis*.

Probabilitas adalah kemungkinan. Hal ini berhubungan dengan terjadinya peristiwa acak. Nilai dinyatakan antara nol dan satu. Probabilitas telah diperkenalkan dalam matematika untuk memprediksi seberapa besar kemungkinan peristiwa akan terjadi. Arti probabilitas pada dasarnya adalah sejauh mana sesuatu mungkin terjadi. Hal tersebut merupakan teori probabilitas dasar yang juga digunakan dalam distribusi probabilitas yaitu tentang kemungkinan hasil untuk eksperimen acak. *Term probability* atau probabilitas kata merupakan salah satu penerapan teori probabilitas. Setiap kata yang terdapat di dalam dokumen merupakan sebuah variabel yang tidak dapat diketahui dengan pasti jumlahnya karena dokumen memiliki banyak kata. Sehingga yang dapat dilakukan hanya

memprediksi kemungkinan sebuah kata terdapat pada dokumen. Perhitungan probabilitas kata sangat bermanfaat untuk menganalisis isi dari sebuah dokumen tanpa membacanya satu per satu.

Teori informasi adalah subbidang matematika yang berkaitan dengan transmisi data. Landasan teori informasi adalah gagasan untuk mengukur seberapa banyak informasi yang ada dalam sebuah pesan. Secara lebih umum, ini dapat digunakan untuk menghitung informasi dalam suatu peristiwa variabel acak, yang disebut *entropy*, dan dihitung menggunakan probabilitas. Pengukuran informasi banyak digunakan dalam *artificial intelligent* seperti membangun *decision tree* dan mengoptimalkan model klasifikasi.

Brownlee (2019) menjelaskan banyaknya informasi yang terdapat dalam variabel acak juga dapat diukur dengan adanya distribusi probabilitas dari variabel acak tersebut. Informasi dalam variabel acak ini disebut dengan *information entropy* atau *Shannon's entropy*. Hal ini terkait dengan gagasan *entropy* dari fisika di mana keduanya berkaitan dengan ketidakpastian.

Menurut Manning (2008), *information retrieval* atau temu kembali informasi adalah pengembalian materi yang tidak terstruktur yang memenuhi kebutuhan informasi dari dalam koleksi yang besar. Dalam penelusuran web, sistem harus menyediakan penelusuran terhadap lebih dari miliaran dokumen yang disimpan dalam jutaan komputer. Prinsip dari temu kembali informasi sangat sederhana. Dokumen-dokumen yang disimpan dalam suatu repositori dapat diambil untuk memenuhi kebutuhan informasi dari *user* yang

mengekspresikannya melalui *query* dan menghasilkan himpunan dokumen yang sesuai dengan *query* tersebut. Terdapat tiga model dari temu kembali informasi, di antaranya adalah *Boolean Model* yang merupakan model sederhana berdasarkan teori himpunan dan aljabar boolean, *Vector Space Model* yang merepresentasikan dokumen dan *query* dalam bentuk vektor dimensional, dan *Probabilistic Model* yang menggunakan teori probabilitas.

BAB III

METODOLOGI PENELITIAN

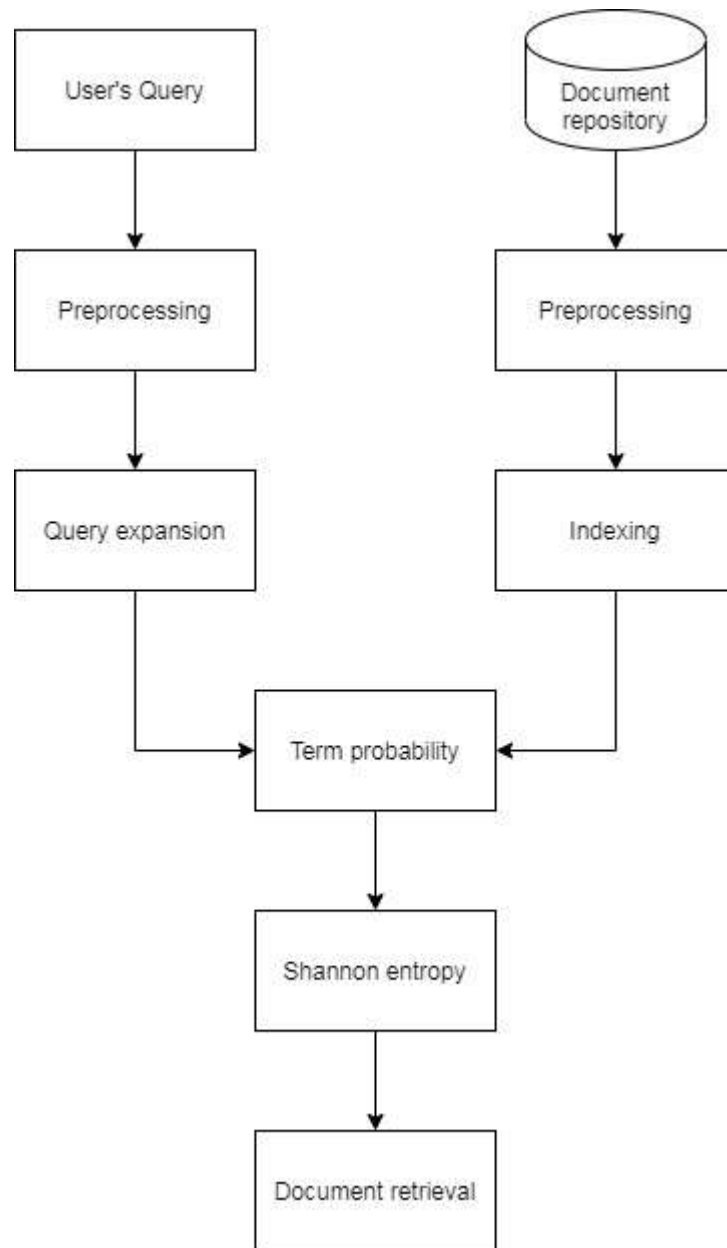
3.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data sekunder berupa dokumen perundang-undangan sejumlah dua ribu dokumen berekstensi *Portable Document Format* (PDF) dengan format teks dari sepuluh bidang kementerian di Indonesia yaitu Kementerian Agama, Kementerian Agraria, Kementerian Kelautan dan Perikanan, Kementerian Pendidikan dan Budaya (Dikbud), Kementerian Kesehatan, Kementerian Keuangan, Kementerian Komunikasi dan Informasi (Kominfo), Kementerian Lingkungan Hidup, Kementerian Pertahanan, serta Kementerian Riset Teknologi dan Pendidikan Tinggi (Ristekdikti) yang didapatkan dari situs – situs di bawah ini :

1. <https://jdih.go.id/>. Situs tersebut merupakan situs Jaringan Dokumentasi dan Informasi Hukum Nasional yang dikelola oleh Badan Pembinaan Hukum Nasional, menyediakan berbagai jenis dokumen hukum dari hukum pusat sampai hukum daerah.
2. <https://peraturan.go.id/>. Situs tersebut merupakan situs yang dikelola oleh Kementerian Hukum dan HAM Republik Indonesia. Kelebihan dari situs tersebut adalah dokumen yang disediakan memiliki format teks yang lebih baik dari situs sebelumnya.

3.2 Desain Sistem

Desain sistem pada penelitian ini dapat dilihat pada Gambar 3.1.

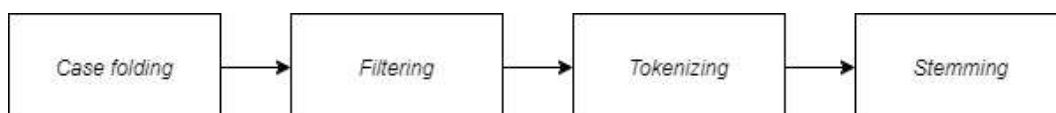


Gambar 3.1. Desain Sistem

3.2.1 Preprocessing

Input pada sistem adalah *query* yang dimasukkan oleh *user*. *Query* berupa kalimat yang akan diproses menjadi kata atau *term* yang digunakan untuk mencari dokumen yang relevan. Batas *query* yang dapat dimasukkan oleh *user* adalah 10

kata setiap satu kali *input*. Selanjutnya, *query* ini melewati tahap *preprocessing* yang terdiri dari *case folding*, *filtering*, *tokenizing*, dan *stemming*. *Preprocessing* juga dilakukan terhadap dokumen yang terdapat dalam repositori. Tahap *preprocessing* dapat dilihat pada Gambar 3.2.



Gambar 3.2. Tahap *Preprocessing*

Case folding merupakan proses konversi huruf menjadi huruf kecil (*lowercase*). Pada proses ini, setiap huruf dari *query* akan diubah menjadi huruf kecil untuk memudahkan identifikasi teks pada proses selanjutnya. Sifat *case sensitive* pada teks membuat komputer membaca huruf kecil dan huruf kapital pada sebuah teks sebagai huruf yang berbeda, hal ini menyebabkan pemrosesan teks menjadi sulit dan tidak efisien. Contoh hasil dari *case folding* dapat dilihat pada Tabel 3.1.

Tabel 3.1 Contoh Hasil *Case Folding*

Sebelum <i>case folding</i>	“Undang-Undang tentang Perguruan Tinggi.”
Setelah <i>case folding</i>	“undang-undang tentang perguruan tinggi.”

Proses selanjutnya adalah *filtering*, yaitu proses untuk menghapus tanda baca, spasi ganda, *new line*, dan *stopwords*. *Stopwords* adalah kata yang tidak memiliki makna sehingga tidak berguna untuk proses identifikasi. Kata-kata tidak bermakna atau tidak penting ini dapat berupa kata hubung (dan, atau, selanjutnya)

atau kata depan (di, ke, pada). Contoh hasil dari proses *filtering* dapat dilihat pada Tabel 3.2.

Tabel 3.2. Contoh Hasil *Filtering*

Sebelum <i>filtering</i>	“undang-undang tentang perguruan tinggi.”
Setelah <i>filtering</i>	“undang undang perguruan tinggi”

Tokenizing adalah proses pemenggalan kalimat menjadi kata-kata penyusunnya atau disebut dengan token. Pemenggalan kalimat didasarkan pada spasi antar kata, dengan kata lain spasi inilah yang menjadi penanda untuk mendeteksi sebuah kata atau bukan. Contoh hasil proses *tokenizing* dapat dilihat pada Tabel 3.3.

Tabel 3.3. Contoh Hasil *Tokenizing*

Sebelum <i>tokenizing</i>	“undang undang perguruan tinggi”
Setelah <i>tokenizing</i>	“undang”, “undang”, “perguruan”, “tinggi”

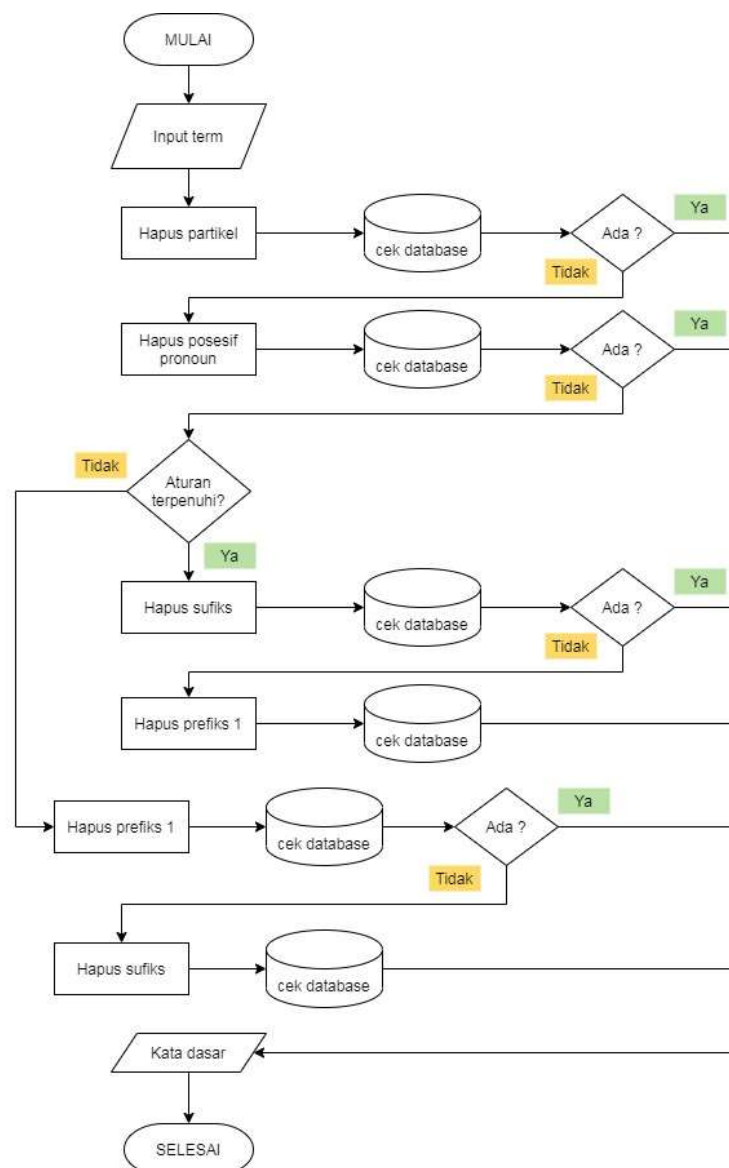
Token yang dihasilkan dari proses *tokenizing* dapat terdiri dari kata dasar dan kata berimbuhan. Dalam Bahasa Indonesia, satu kata dasar dapat menghasilkan banyak kata lain jika memiliki imbuhan baik awalan atau akhiran. Maka dari itu, kata berimbuhan harus diubah menjadi kata dasar untuk mempermudah sistem dalam mengidentifikasi kata sehingga pemrosesan kata menjadi lebih efisien. Proses ini disebut dengan *stemming*. Algoritma *stemming* yang digunakan dalam penelitian ini adalah algoritma Porter yang telah dikembangkan oleh Tala (2003) untuk memproses teks Bahasa Indonesia.

Contoh hasil proses *stemming* dapat dilihat pada Tabel 3.4.

Tabel 3.4. Contoh Hasil *Stemming*

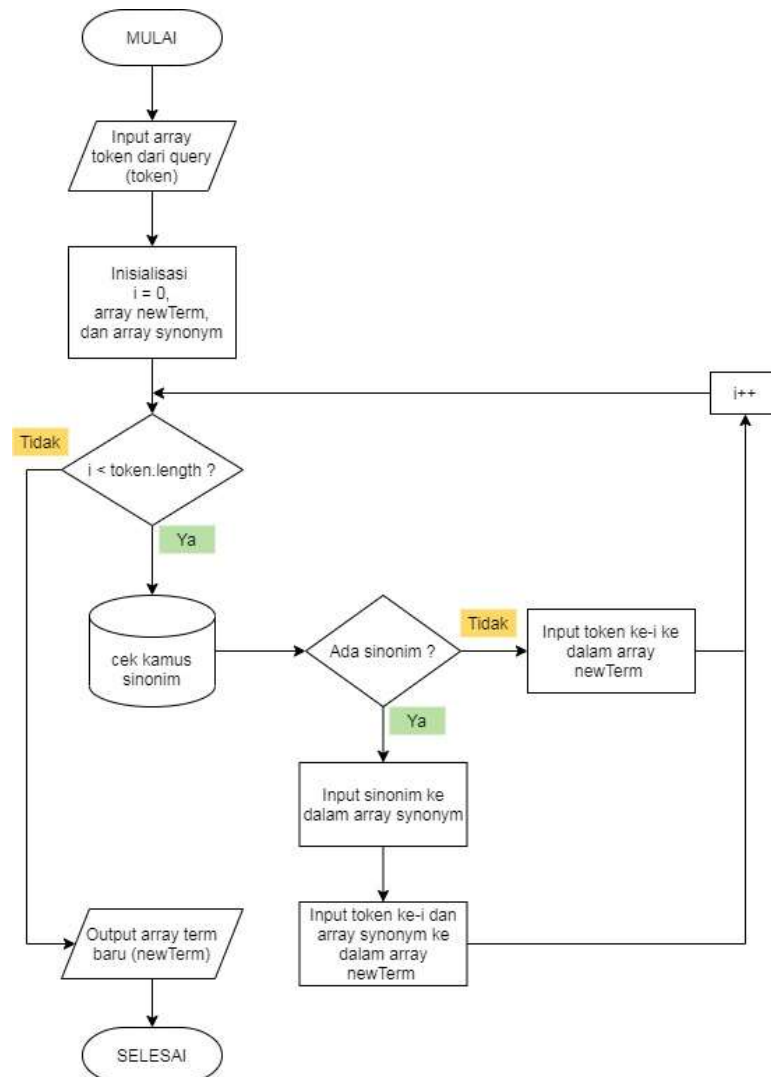
Kata	Imbuhan	Pengganti	Hasil <i>stemming</i>
Menimbang	men- (awalan)	t	Timbang
Memotong	mem- (awalan)	p	Potong
Penyakit	peny- (awalan)	s	Sakit
Layanan	-an (akhiran)	-	Layan
Peraturan	per- (awalan) -an (akhiran)	-	Atur

Proses *stemming* dapat dilihat pada Gambar 3.3 di bawah ini.

Gambar 3.3. *Flowchart Stemming*

3.2.2 Query Expansion

Dokumen yang dikembalikan berdasarkan *query* dapat diperluas dengan cara memperluas *query*. Perluasan *query* dapat dilakukan dengan cara menyusun ulang *query* dengan menambahkan kata yang bermakna sama (sinonim) dengan kata yang terdapat pada *query* awal. Sinonim diambil dari kamus yang telah dibangun sebelumnya menggunakan kumpulan sinonim kata yang diambil dari Tesaurus Bahasa Indonesia. Proses *query expansion* dapat dilihat pada Gambar 3.4.



Gambar 3.4. Flowchart Query Expansion

Kamus sinonim dibangun dengan cara mengambil dua sampai tiga kata sinonim dari Tesaurus Bahasa Indonesia untuk kata tertentu. Contoh kamus sinonim yang dibangun dapat dilihat pada Tabel 3.5.

Tabel 3.5. Contoh Kamus Sinonim

Kata	Sinonim
Agama	Ajaran, kepercayaan, keyakinan
Didik	Ajar, asuh, bimbing,
Pajak	Cukai, retribusi, tariff
Warga	Penduduk, penghuni

Setelah kamus sinonim dibangun, setiap kata pada *query* akan dicek untuk menemukan sinonim yang terdapat pada kamus. Apabila kata pada *query* ditemukan pada kamus, maka gugus sinonim dari kata tersebut beserta kata itu sendiri akan dikembalikan untuk menyusun *query* baru. Apabila kata pada *query* tidak ditemukan, maka kata tersebut dikembalikan untuk menyusun *query* baru tanpa sinonimnya. Contoh hasil perluasan *query* dapat dilihat pada Tabel 3.6.

Tabel 3.6. Contoh *Query Expansion*

<i>Query Awal</i>	<i>Preprocessing</i>	<i>Hasil Query Expansion</i>
Pendidikan warga	didik warga	didik ajar asuh bimbing warga penduduk penghuni

Setelah dilakukan *query expansion*, kata yang terdapat pada *query* yang baru akan digunakan pada proses selanjutnya untuk menemukan dokumen yang relevan dengan cakupan yang lebih luas.

Implementasi dari algoritma *query expansion* dapat dilihat pada Gambar 3.5 di bawah ini.

```
# QUERY EXPANSION
def queryPreprocess(text):
    queryAwal = preprocessing.preprocess(text)
    sinonim = []
    term = []
    kata = synonym.objects.all().values_list('kata', flat=True)
    sin = synonym.objects.all().values_list('sinonim', flat=True)

    for x in sin:
        s = list(x.split(' '))
        sinonim.append(s)

    for x in kata:
        term.append(x)

    word = dict(zip(term, sinonim))

    a = ' '
    q = []
    for x in queryAwal:
        q.append(x)
        for y in word:
            if x == y:
                for z in word[y]:
                    q.append(z)

    queryExpansion = a.join(q)
    queryAkhir = preprocessing.preprocess(queryExpansion)

    return queryAkhir
```

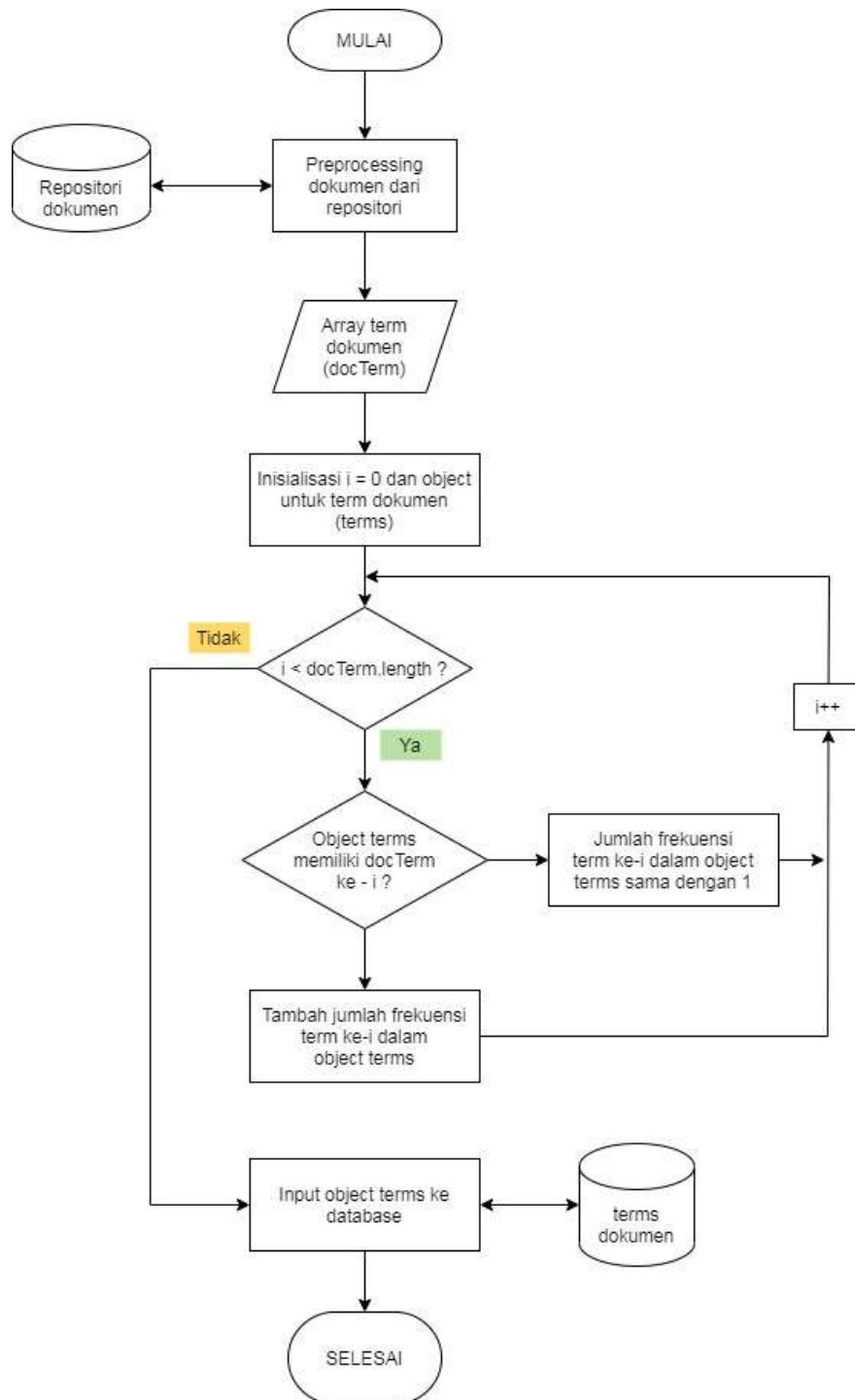
Gambar 3.5. Source Code Query Expansion

Penjelasan dari *source code* di atas adalah *query* awal akan melewati proses *preprocessing*, setelah itu dilakukan proses pencarian sinonim terhadap kata yang sama di dalam database. Sinonim yang ditemukan di dalam database

akan ditambahkan ke dalam gugus *term* hasil *preprocessing* pada *query* awal dan menjadi gugus *term* yang baru. *Preprocessing* diterapkan lagi setelah *query expansion* untuk mendapatkan kata dasar dari sinonim yang masih berimbuhan, sehingga gugus *term* hasil akhir dapat digunakan pada proses selanjutnya.

3.2.3 Indexing

Dokumen yang telah melalui proses *preprocessing* akan menghasilkan kumpulan *term* yang menyusun dokumen tersebut. Selanjutnya setiap *term* akan dihitung jumlah kemunculannya di dalam dokumen. Hasil perhitungan masing-masing *term* inilah yang disebut dengan *term frequency* yang akan disimpan di dalam *database* untuk proses selanjutnya yaitu perhitungan *term probability* yang bersesuaian dengan *query*. Proses penyimpanan *term* dan jumlah di dalam dokumen inilah yang disebut dengan *indexing*, sehingga sistem tidak perlu menjalankan *preprocessing* untuk dokumen lagi setiap kali *query* dimasukkan oleh *user* karena dapat memperlambat kinerja sistem. Proses perhitungan *term frequency* dapat dilihat pada Gambar 3.6.



Gambar 3.6. *Flowchart Indexing*

Implementasi dari proses *indexing* dapat dilihat pada Gambar 3.7 di bawah ini.

```
def inDB(names, terms, freq):
    hasil = termIndex(
        nama_doc = names,
        term = terms,
        totalFreq = freq
    )
    hasil.save()
    print('{} successfully saved.'.format(names))

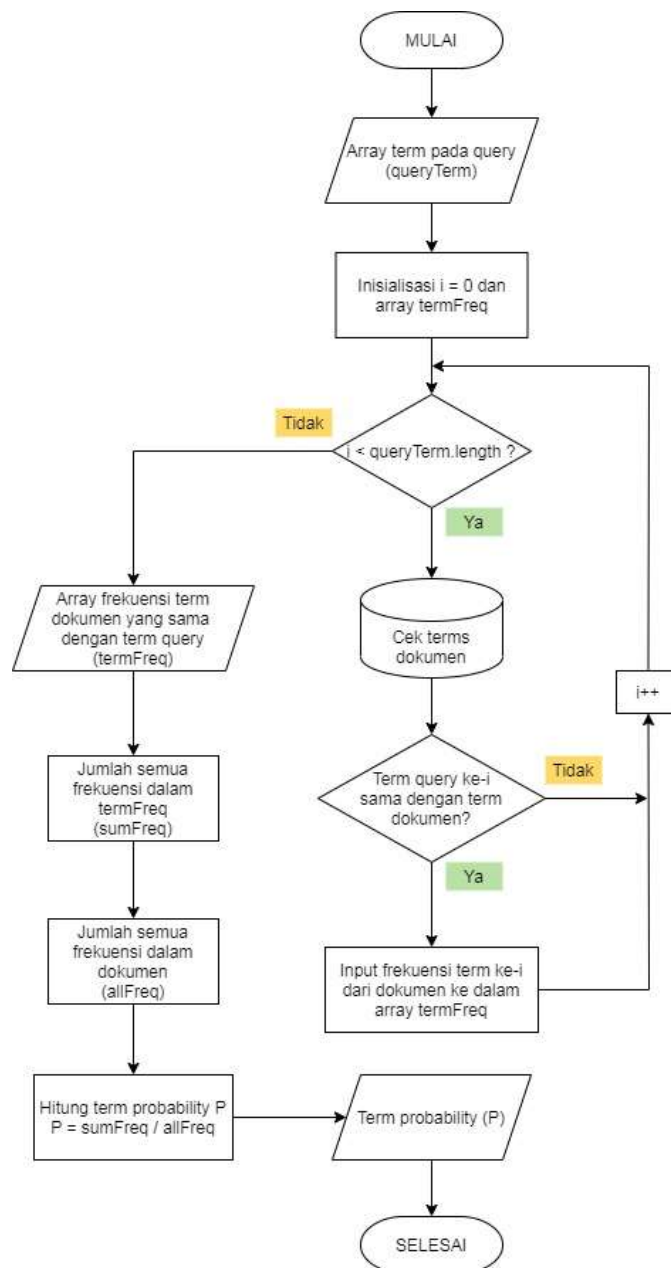
def main():
    for key in range(len(pdfVal)):
        name = pdfName[key]
        hasil = docPreprocess(pdfVal[key])
        term = frequencyDoc(hasil)
        allFre = frequencyAll(hasil)
        cek = termIndex.objects.filter(nama_doc=name)
        if cek.exists():
            print('{} already in database.'.format(name))
            break
        else:
            db = inDB(name, term, allFre)
    return name
```

Gambar 3.7. *Source Code Indexing*

Berdasarkan *source code* di atas, atribut dari dokumen yaitu nama, *terms*, dan total *term frequency* dimasukkan ke dalam database. Proses *indexing* dilakukan secara urut untuk menghindari duplikasi dokumen pada database.

3.2.4 *Term Probability*

Alur mendapatkan *term probability* dapat dilihat pada Gambar 3.8 berikut.



Gambar 3.8. Flowchart Term Probability

Term probability didapatkan dari menghitung perbandingan *term frequency* yang terdapat pada dokumen terhadap *query* dengan total frekuensi *term* yang terdapat pada dokumen tersebut. *Frequency term* pada proses ini didapatkan dengan cara mencari *term* pada dokumen yang bersesuaian dengan *term* pada

query di dalam *database* yang telah dibangun pada proses *indexing*. Sedangkan total frekuensi *term* didapatkan dari jumlah seluruh *term* atau kata pada dokumen.

Setelah mendapatkan frekuensi *term*, maka *term probability* dapat ditentukan dengan persamaan

$$\text{term probability} = \frac{\text{term frequency}}{\text{total term frequency}} \quad (3.1)$$

Implementasi algoritma dari *term probability* dapat dilihat pada Gambar 3.9 di bawah ini.

```
def similarity():
    kata = []
    freq = []
    tokenDoc = takeDB()
    query = list(set(queryPreprocess()))
    # term yg sama dengan query
    for x in range(len(tokenDoc)):
        kata.append([])
        for y in tokenDoc[x].keys():
            for q in query:
                if q == y:
                    kata[x].append(y)
    # ambil frekuensi term yg sama dengan query
    for k in range(len(tokenDoc)):
        freq.append([])
        for m in kata[k]:
            if m in tokenDoc[k]:
                freq[k].append(tokenDoc[k][m])
    # ambil total frekuensi dari db
    total = []
    prob = []
    totalFreq = termIndex.objects.all().values_list('totalFreq', flat=True)
    for n in totalFreq:
        total.append(n)
    # probability
    for x in range(len(freq)):
        prob.append([])
        for y in freq[x]:
            p = y/total[x]
            prob[x].append(p)
```

Gambar 3.9 Source Code Term Probability

Pada *source code* di atas, proses pertama dalam *term probability* adalah mencari *term* pada dokumen di database yang sama dengan *query*. Setelah itu didapatkan *frequency* dari *term* tersebut. Untuk dapat masuk ke dalam rumus *probability*, sistem membutuhkan total *term frequency* dari dokumen yang didapat pada proses sebelumnya. Maka sistem akan mengambil total *term frequency* dari database dan menghitung *term probability* sesuai rumus.

3.2.5 *Shannon's Entropy*

Penelitian ini menggunakan metode *Shannon's entropy* dengan alasan bahwa *Shannon's entropy* merupakan cara kuantitatif dalam mengukur keberagaman *term* dalam sebuah dokumen. Metode ini cocok untuk mengukur eksistensi *term* dalam dokumen yang sama dengan *query user* sehingga dapat mendeteksi dokumen yang relevan dengan kebutuhan informasi *user*.

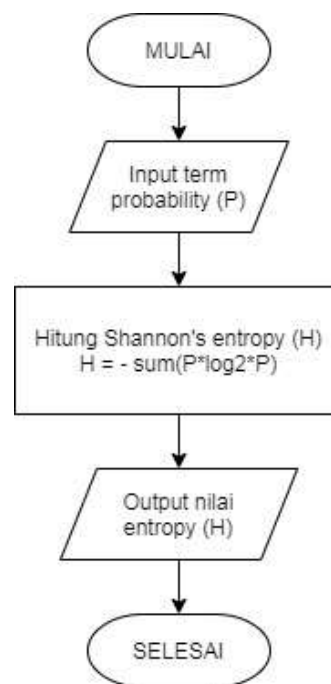
Shannon's entropy menghitung rata-rata informasi dari kumpulan suatu kejadian. Semakin tinggi nilai *entropy* yang didapatkan, maka semakin tinggi ketidakpastian atau informasi yang terkandung dalam sebuah kejadian. Sebaliknya, semakin rendah nilai *entropy* yang didapatkan, maka semakin sedikit pula ketidakpastian atau informasi yang terkandung dalam sebuah kejadian. Perhitungan *Shannon's entropy* ditunjukkan dengan rumus berikut

$$H(x) = E_x[I(x)] = - \sum_{x=1}^X P(x) \ln P(x) \quad (3.2)$$

Notasi *ln* dalam rumus *entropy* memiliki arti *log* basis 2 yang berarti informasi diukur dalam satuan bit atau digit biner. Semakin besar nilai *entropy*

suatu dokumen maka dokumen tersebut akan semakin relevan dengan *query user*. Sebaliknya, apabila suatu dokumen memiliki nilai *entropy* rendah maka dokumen tersebut memiliki relevansi kecil terhadap *query user*. Gugus dokumen relevan yang didapatkan dari perhitungan *entropy* akan diranking berdasarkan nilai *entropy* tertinggi (*maximum*) sampai terendah untuk mendapatkan dokumen paling relevan di antara dokumen relevan lainnya.

Flowchart tahapan mendapatkan nilai *entropy* dapat dilihat dari Gambar 3.10 di bawah ini.



Gambar 3.10 *Flowchart Entropy*

Gambar 3.11 menunjukkan mengenai implementasi algoritma *Shannon's Entropy* melalui *source code*.

```

# entropy
information = []
for x in range(len(prob)):
    information.append([])
    for y in prob[x]:
        en = -log2(y)*y
        information[x].append(en)
entropy = []
for x in range(len(information)):
    entropy.append([])
    en = sum(information[x])
    entropy[x].append(en)

return entropy

```

Gambar 3.11. *Source Code Entropy*

Variabel *information* pada *source code* di atas digunakan untuk menyimpan nilai $-\log$ basis 2 dari *term probability* pada dokumen. Setelah itu setiap *information* akan dijumlah untuk mendapatkan nilai *entropy* setiap dokumen. Hasil dari pengembalian dokumen akan diranking sejumlah 10 teratas berdasarkan nilai *entropy* yang didapatkan. *Source code* perangkangan dapat dilihat pada Gambar 3.12 di bawah ini.

```

# ranking
def rank():
    finalResult = []
    name = totalFreq = termIndex.objects.all().values_list('nama_doc', flat=True)
    ent = similarity()
    dict_entropy = dict(zip(name, ent))
    sort = {sort : v for sort, v in sorted(dict_entropy.items(), reverse=True, key=lambda item: item[1])}

    arr = list(sort)
    for x in range(10):
        finalResult.append(arr[x])

    return sort

```

Gambar 3.12. *Source Code Ranking*

BAB IV

UJI COBA DAN PEMBAHASAN

Pada bab ini akan dibahas tentang implementasi dari langkah-langkah pada skenario uji coba yang telah dibahas pada bab sebelumnya. Setelah itu akan ditunjukkan hasil dan pembahasan dari uji coba, serta integrasi penelitian ini dengan islam.

4.1 Langkah Uji Coba

Pengujian yang dilakukan pada sistem ini dilakukan untuk mengukur *accuracy*, *precision*, *recall*, dan *f-measure* dari metode yang digunakan yaitu metode *Shannon's entropy* dengan *query expansion*.

4.1.1 *Preprocessing*

Preprocessing dilakukan terhadap dokumen dan *query* yang dimasukkan oleh pengguna. Tahap *preprocessing* terdiri dari *case folding*, *filtering*, *tokenizing*, dan *stemming*. *Preprocessing* yang dilakukan terhadap *query* pengguna akan menghasilkan kata tunggal atau *term*. Contoh hasil dari *preprocessing query* dapat dilihat pada Tabel 4.1 di bawah ini.

Tabel 4.1. Contoh Hasil *Preprocessing Query*

	Contoh <i>Query</i>	
<i>Query awal</i>	Universitas dan perguruan tinggi	Jaminan kesehatan
Hasil <i>preprocessing</i>	universitas, guru, tinggi	jamin, sehat

Preprocessing dokumen dilakukan terhadap dua ribu dokumen yang berekstensi PDF. Contoh hasil dari *preprocessing* dokumen dapat dilihat pada Tabel 4.2 di bawah.

Tabel 4.2. Contoh Hasil *Preprocessing* Dokumen

	Contoh Penggalan Dokumen
Kalimat awal	PERATURAN PEMERINTAH REPUBLIK INDONESIA NOMOR 49 TAHUN 2005 TENTANG PEDOMAN KEGIATAN PELIPUTAN LEMBAGA PENYIARAN ASING DENGAN RAHMAT TUHAN YANG MAHA ESA PRESIDEN REPUBLIK INDONESIA,
Hasil <i>preprocessing</i>	'atur', 'perintah', 'republik', 'indonesia', 'nomor', 'pedoman', 'giat', 'liput', 'lembaga', 'siar', 'asing', 'rahmat', 'tuhan', 'maha', 'esa', 'presiden', 'republik', 'indonesia'

4.1.2 Query Expansion

Setelah *preprocessing*, dilakukan *query expansion* pada *query* yang dimasukkan oleh pengguna. Proses ini berfungsi untuk menambahkan *term* lain yang berkaitan dengan kata kunci. *Term* yang berkaitan didapatkan dari kamus sinonim yang telah dibangun sebelumnya menggunakan Tesaurus. Contoh daftar sinonim pada kamus yang telah dibangun dapat dilihat pada Tabel 4.3 di bawah ini.

Tabel 4.3. Contoh Kamus Sinonim

No	Term	Sinonim
1	adat	budaya tradisi
2	manfaat	faedah fungsi
3	dana	anggaran biaya modal
4	didik	asuh bimbing ajar
5	fasilitas	alat sarana layanan
6	hutan	alas rimba
7	tanah	lahan
8	universitas	perguruan tinggi institut akademi

9	wilayah	daerah kawasan lingkungan
10	frekuensi	gelombang saluran

Langkah dalam proses ini adalah dengan mengecek satu persatu kata yang terdapat dalam *query*, apabila kata tersebut terdapat dalam kamus sinonim, maka gugus sinonim akan diambil dan ditambahkan ke *query* awal sehingga menjadi *query* baru yang memiliki lebih banyak *term* yang saling berkaitan.

Contoh hasil perluasan *query* dapat dilihat pada Tabel 4.4 di bawah ini.

Tabel 4.4. Contoh Hasil *Query Expansion*

	<i>Query</i>
<i>Query</i> awal	Pemanfaatan hutan
Hasil <i>preprocessing</i>	manfaat, hutan
Hasil <i>query expansion</i>	manfaat, faedah, fungsi, hutan, alas, rimba

Pada kamus sinonim yang telah dibangun, kata ‘manfaat’ memiliki sinonim ‘faedah’ dan ‘fungsi’, kata ‘hutan’ memiliki sinonim ‘alas’ dan ‘rimba’ sehingga *query* awal hasil *preprocessing* akan disusun ulang menjadi ‘manfaat faedah fungsi hutan alas rimba’.

4.1.3 *Indexing*

Tahap selanjutnya yang dilakukan pada dokumen adalah *indexing*, yaitu menghitung frekuensi dari tiap *term* dan frekuensi seluruh *term* yang terdapat pada dokumen. Frekuensi yang dimaksud adalah *raw frequency* yaitu jumlah kemunculan *term* pada sebuah dokumen. Hasil dari *indexing* disimpan ke dalam

DBMS MySQL. Contoh tabel yang digunakan untuk menampung hasil *indexing* dapat dilihat pada Tabel 4.5 di bawah.

Tabel 4.5. Contoh Hasil *Indexing*

id	nama_doc	term	totalFreq
1	PERPRES No. 41 Tahun 2010.pdf	{'presiden': 18, 'republik': 13, 'indonesia': 13, 'atur': 13, 'nomor': 11, 'bija': 12, 'tahan': 20, 'negara': 25, 'rahmat': 1, 'tuhan': 1, 'maha': 1, 'esa': 1, 'timbang': 2, 'salah': 1, 'fungsi': 2, 'perintah': 2, 'laksana': 2, 'sistem': 4, 'rakyat': 1, 'semesta': 1, ...}	278
2	PP No. 11 Tahun 2003.pdf	{'atur': 17, 'perintah': 14, 'republik': 5, 'indonesia': 5, 'nomor': 19, 'rubah': 2, 'gaji': 6, 'gawai': 7, 'neger': 7, 'sipil': 7, 'kali': 4, 'ubah': 5, 'presiden': 2, 'timbang': 1, 'rangka': 1, 'jamin': 1, 'penuh': 1, 'butuh': 1, 'dasar': 2, 'terima': 1, 'tingkat': 1, 'daya': 2, 'hasil': 1, 'sejahtera': 1, ...}	179
3	UU No. 2 Tahun 1947.pdf	{'presiden': 5, 'republik': 5, 'indonesia': 5, 'undang': 12, 'nomor': 1, 'dewan': 7, 'tahan': 7, 'negara': 9, 'atur': 13, 'pengesyahan': 1, 'no': 27, 'timbang': 1, 'kuasa': 1, 'sahkan': 1, 'pasal': 5, 'bahaya': 2, 'ayat': 1, 'iv': 1, 'alih': 1, 'dasar': 1, 'maklumat': 1, 'wakil': 1, 'tanggal': 14, 'oktober': 2, 'persetujuan': 1, 'badan': 2, 'kerja': 1, ...}	181

Penyimpanan *term* pada *database* bertujuan agar proses selanjutnya tidak memerlukan *preprocessing* dokumen terlebih dahulu sehingga mempersingkat waktu pengembalian dokumen.

4.1.4 *Term Probability*

Langka selanjutnya yaitu menghitung *term probability* yang terdapat pada *query* yang telah diperluas pada proses *query expansion*. Sistem akan mencari *term* yang sama antara *query* dan dokumen, lalu mengambil *frequency term*

tersebut dan membandingkannya dengan total *term* pada dokumen. Tabel 4.6 di bawah ini merupakan contoh hasil perhitungan *term probability* dengan *query* “pemanfaatan hutan” dan hasil *query expansion* “manfaat fungsi faedah hutan alas rimba”.

Tabel 4.6. Contoh Hasil Perhitungan *Term Probability*

No	Dokumen	<i>Term</i> yang sama	<i>Probability</i>
1	PM Lingkungan Hidup No. 37 Tahun 2019.pdf	fungsi	0,0061
2	PM Lingkungan Hidup No. 69 Tahun 2019.pdf	manfaat	0,0367
		hutan	0,1063
3	PM Kehutanan No. 47 Tahun 2013.pdf	fungsi	0,0019
		hutan	0,0871
		manfaat	0,0407
4	PP No. 34 Tahun 2002.pdf	fungsi	0,0040
		hutan	0,0911
		manfaat	0,0322
		fungsi	0,0029
5	UU No. 41 Tahun 1999.pdf	rimba	0,0002
		alas	0,0011
		hutan	0,0945
		manfaat	0,0180
		fungsi	0,0090
6	PP No. 10 Tahun 2010.pdf	rimba	0,0002
		fungsi	0,0238
		hutan	0,0979
7	PM Lingkungan Hidup No. 45 Tahun 2016.pdf	manfaat	0,0037
		hutan	0,0776
		manfaat	0,0265
		fungsi	0,0046
8	PP No. 6 Tahun 2007.pdf	alas	0,0027
		hutan	0,0946
		manfaat	0,0235
		fungsi	0,0036
		rimba	0,0079
9	PM Lingkungan Hidup No. 75 Tahun 2016.pdf	alas	0,0003
		hutan	0,0591
		manfaat	0,0039
10	PM Lingkungan Hidup No. 49 Tahun 2017.pdf	hutan	0,0649
		manfaat	0,0337
		fungsi	0,0039

Tidak semua dokumen memiliki *term* yang sama dengan *query*. Dari seluruh dokumen yang terdapat pada repositori, *term probability* dari dokumen tersebut akan bernilai 0 sehingga tidak akan dikembalikan.

4.1.5 *Shannon's Entropy*

Setelah dihitung *term probability* dari setiap dokumen, selanjutnya adalah menghitung *entropy* dari setiap dokumen tersebut. Proses ini merupakan proses akhir yang menentukan dokumen-dokumen yang akan dikembalikan. Nilai *entropy* didapatkan dengan cara menjumlahkan *term probability* pada proses sebelumnya dan mensubstitusikannya ke rumus 3.3. Contoh hasil perhitungan *entropy* dengan query “Pemanfaatan hutan” dapat dilihat pada Tabel 4.7 di bawah ini.

Tabel 4.7. Contoh Hasil Perhitungan *Entropy*

No	Dokumen	<i>Entropy</i>
1	PM Lingkungan Hidup No. 37 Tahun 2019.pdf	0,6037
2	PM Lingkungan Hidup No. 69 Tahun 2019.pdf	0,5362
3	PM Kehutanan No. 47 Tahun 2013.pdf	0,5265
4	PP No. 34 Tahun 2002.pdf	0,5127
5	UU No. 41 Tahun 1999.pdf	0,4902
6	PP No. 10 Tahun 2010.pdf	0,4862
7	PM Lingkungan Hidup No. 45 Tahun 2016.pdf	0,4837
8	PP No. 6 Tahun 2007.pdf	0,4826
9	PM Lingkungan Hidup No. 75 Tahun 2016.pdf	0,4729
10	PM Lingkungan Hidup No. 49 Tahun 2017.pdf	0,4523

Dari perhitungan nilai *entropy*, dokumen akan diranking sejumlah 10 teratas untuk dikembalikan. *Entropy* menunjukkan seberapa relevan dokumen dengan *query* yang diberikan sehingga semakin besar nilai *entropy* maka semakin relevan pula dokumen tersebut.

4.2 Hasil Uji Coba

Uji coba dilakukan dua kali menggunakan 30 *query* berdasarkan dokumen yang terdapat pada repositori. Berdasarkan *query* tersebut, dapat diukur akurasi, presisi, *recall*, dan *f-measure* sistem dengan menggunakan *Shannon's entropy* tanpa *query expansion* dan menggunakan *query expansion*. Daftar *query* uji coba dapat dilihat pada Tabel 4.8 di bawah ini.

Tabel 4.8. Daftar *Query* Uji Coba

No	<i>Query</i>
1	Tunjangan hari raya untuk pegawai
2	Pemanfaatan hutan
3	Penyelenggaraan penyiaran
4	Jaringan telekomunikasi berbasis internet
5	Persyaratan perizinan jasa penyiaran televisi
6	Ibadah Haji
7	Manfaat hutan untuk menanggulangi banjir dan erosi
8	Perkawinan
9	Fungsi pertahanan bagi sebuah negara
10	Gaji pegawai
11	Persyaratan penanaman modal usaha
12	Rehabilitasi wilayah pesisir
13	Misi pemeliharaan perdamaian
14	Penyelenggaraan perjalanan ibadah umrah
15	Larangan mengeluarkan benih ikan dari indonesia
16	Organisasi dan tata kerja balai penelitian
17	Ekspor dan impor narkotika
18	Pengadaan vaksin corona
19	Pendirian rumah sakit umum
20	Bantuan kesehatan dari pemerintah
21	Budidaya perikanan
22	Tarif layanan badan layanan umum
23	Anggaran pendapatan dan belanja negara
24	Pengaturan penggunaan frekuensi radio
25	Pembiayaan telekomunikasi di indonesia
26	Tarif sewa saluran siaran
27	Universitas dan perguruan tinggi
28	Jaminan kesehatan
29	Organisasi dan tata kerja balai pengelolaan hutan

30	Tata cara pembuangan limbah ke laut
----	-------------------------------------

Data 10 teratas yang digunakan sebagai *groundtruth* atau acuan relevanan dokumen yang diambil dari situs <https://www.google.com/> berdasarkan 30 *query* uji coba dapat dilihat pada Lampiran 1.

Dokumen yang dikembalikan akan dirangking berdasarkan nilai *entropy* sebanyak 10 teratas untuk kemudian ditampilkan. Jika 10 dokumen teratas yang terdapat pada hasil sistem ditemukan pula di lingkup 10 teratas dokumen *grountruth* dengan *query* yang sama, maka dokumen tersebut dikatakan relevan. Apabila 10 dokumen teratas yang terdapat pada hasil sistem tidak ditemukan di lingkup 10 teratas dokumen *groundtruth*, maka dokumen tersebut dikatakan tidak relevan. Nilai *entropy* yang didapatkan memiliki 4 angka di belakang koma, hal ini bertujuan untuk mendapatkan nilai *entropy* yang lebih detail sehingga sistem dapat melakukan perankingan dokumen dengan baik. Daftar hasil pengembalian dokumen tanpa menggunakan *query expansion* dapat dilihat secara lengkap pada Lampiran 2. Hasil uji coba tanpa menggunakan *query expansion* ditunjukkan pada Tabel 4.9 di bawah ini.

Tabel 4.9. Hasil Uji Coba Tanpa Menggunakan *Query Expansion*

No	<i>Query</i>	TP	FP	FN	TN
1	Tunjangan hari raya untuk pegawai	5	5	5	1985
2	Pemanfaatan hutan	4	6	6	1984
3	Penyelenggaraan penyiaran	4	6	6	1984
4	Jaringan telekomunikasi berbasis internet	5	5	5	1985
5	Persyaratan perizinan jasa penyiaran televisi	4	6	6	1984
6	Ibadah Haji	6	4	4	1986
7	Manfaat hutan untuk menanggulangi banjir dan erosi	3	7	7	1983
8	Perkawinan	8	2	2	1988
9	Fungsi pertahanan bagi sebuah negara	2	8	8	1982
10	Gaji pegawai	7	3	3	1987

11	Persyaratan penanaman modal usaha		7	3	3	1987
12	Rehabilitasi wilayah pesisir		3	7	7	1983
13	Misi pemeliharaan perdamaian		7	3	3	1987
14	Penyelenggaraan perjalanan ibadah umrah		7	3	3	1987
15	Larangan mengeluarkan benih ikan dari indonesia		3	7	7	1983
16	Organisasi dan tata kerja balai penelitian		3	7	7	1983
17	Ekspor dan impor narkoba		5	5	5	1985
18	Pengadaan vaksin corona		5	5	5	1985
19	Pendirian rumah sakit umum		4	6	6	1984
20	Bantuan kesehatan dari pemerintah		1	9	9	1981
21	Budidaya perikanan		3	7	7	1983
22	Tarif layanan badan layanan umum		0	10	10	1980
23	Anggaran pendapatan dan belanja negara		2	8	8	1982
24	Pengaturan penggunaan frekuensi radio		5	5	5	1985
25	Pembiayaan telekomunikasi di indonesia		1	9	9	1981
26	Tarif sewa saluran siaran		4	6	6	1984
27	Universitas dan perguruan tinggi		2	8	8	1982
28	Jaminan kesehatan		4	6	6	1984
29	Organisasi dan tata kerja balai pengelolaan hutan		4	6	6	1984
30	Tata cara pembuangan limbah ke laut		6	4	4	1986
Total			124	176	176	59524

Uji coba kedua menggunakan *query* yang telah diperluas. Daftar hasil pengembalian dokumen menggunakan *query expansion* dapat dilihat secara lengkap pada Lampiran 3. Hasil uji coba menggunakan *query expansion* ditunjukkan pada Tabel 4.10 di bawah ini.

Tabel 4.10. Hasil Uji Coba Menggunakan *Query Expansion*

No	Query	TP	FP	FN	TN
1	Tunjangan hari raya untuk pegawai	5	5	5	1985
2	Pemanfaatan hutan	5	5	5	1985
3	Penyelenggaraan penyiaran	5	5	5	1985
4	Jaringan telekomunikasi berbasis internet	5	5	5	1985
5	Persyaratan perizinan jasa penyiaran televisi	5	5	5	1985
6	Ibadah Haji	6	4	4	1986
7	Manfaat hutan untuk menanggulangi banjir dan erosi	3	7	7	1983
8	Perkawinan	8	2	2	1988
9	Fungsi pertahanan bagi sebuah negara	2	8	8	1982
10	Gaji pegawai	7	3	3	1987

11	Persyaratan penanaman modal usaha	7	3	3	1987
12	Rehabilitasi wilayah pesisir	4	6	6	1984
13	Misi pemeliharaan perdamaian	7	3	3	1987
14	Penyelenggaraan perjalanan ibadah umrah	7	3	3	1987
15	Larangan mengeluarkan benih ikan dari indonesia	2	8	8	1982
16	Organisasi dan tata kerja balai penelitian	3	7	7	1983
17	Ekspor dan impor narkoba	4	6	6	1984
18	Pengadaan vaksin corona	5	5	5	1985
19	Pendirian rumah sakit umum	4	6	6	1984
20	Bantuan kesehatan dari pemerintah	2	8	8	1982
21	Budidaya perikanan	4	6	6	1984
22	Tarif layanan badan layanan umum	0	10	10	1980
23	Anggaran pendapatan dan belanja negara	2	8	8	1982
24	Pengaturan penggunaan frekuensi radio	5	5	5	1985
25	Pembiayaan telekomunikasi di indonesia	1	9	9	1981
26	Tarif sewa saluran siaran	4	6	6	1984
27	Universitas dan perguruan tinggi	4	6	6	1984
28	Jaminan kesehatan	5	5	5	1985
29	Organisasi dan tata kerja balai pengelolaan hutan	4	6	6	1984
30	Tata cara pembuangan limbah ke laut	6	4	4	1986
Total		131	169	169	59531

Dari Tabel 4.9, dapat diukur hasil evaluasi sistem pengembalian dokumen perundang-undangan menggunakan metode *Shannon's entropy* tanpa menggunakan *query expansion* sebagai berikut

$$Precision = \frac{124}{124+176} \times 100 \% = 41,33 \% \quad (4.1)$$

$$Recall = \frac{124}{176+176} \times 100 \% = 41,33 \% \quad (4.2)$$

$$F - measure = 2 \times \frac{(41,33 \times 41,33)}{41,33+41,33} \times 100 \% = 41,33 \% \quad (4.3)$$

$$Accuracy = \frac{124+59524}{60000} \times 100 \% = 99,41 \% \quad (4.4)$$

Sedangkan dari hasil uji coba sistem pengembalian dokumen perundang-undangan menggunakan metode *Shannon's entropy* dan *query expansion* pada Tabel 4.10, dapat diukur hasil evaluasi dan *query expansion* sebagai berikut

$$Precision = \frac{131}{131+169} \times 100 \% = 43,67 \% \quad (4.5)$$

$$Recall = \frac{131}{131+169} \times 100 \% = 43,67 \% \quad (4.6)$$

$$F - measure = 2 \times \frac{(43,67 \times 43,67)}{43,67+43,67} \times 100 \% = 43,67 \% \quad (4.7)$$

$$Accuracy = \frac{131+59531}{60000} \times 100 \% = 99,44 \% \quad (4.8)$$

Maka hasil pengukuran yang didapatkan tanpa menggunakan *query expansion* adalah nilai *precision* sebesar 41,33%, nilai *recall* sebesar 41,33%, nilai *f-measure* sebesar 41,33% dan nilai *accuracy* sebesar 99,41%. Sedangkan hasil pengukuran yang didapatkan menggunakan *query expansion* adalah nilai *precision* sebesar 43,67%, nilai *recall* sebesar 43,67%, nilai *f-measure* sebesar 43,67% dan nilai *accuracy* sebesar 99,44% dengan peningkatan nilai *precision*, *recall*, dan *f-measure* sebesar 2,34% dan nilai *accuracy* sebesar 0,3%.

4.3 Pembahasan

Berdasarkan hasil uji coba yang telah didapatkan, diketahui bahwa metode *Shannon's entropy* dapat mengembalikan dokumen berdasarkan tingkat korelevanannya terhadap *query* yang diberikan. Tingkat korelevanan sangat bergantung pada jumlah *term* yang terdapat pada dokumen atau yang disebut dengan *term probability*. Semakin besar nilai *probability* yang dimiliki oleh

sebuah dokumen, maka semakin besar pula nilai *entropy* yang dihasilkan. Sedangkan metode *query expansion* dapat membuat sebuah dokumen memiliki nilai *probability* yang semakin tinggi. *Query* yang diperluas akan memiliki *term* yang berkaitan dengan *query* awal menjadi lebih banyak sehingga nilai *probability* pada dokumen akan menjadi lebih besar dan peluang untuk dikembalikan juga semakin besar.

Berdasarkan pembahasan di atas, sistem dapat mengembalikan dokumen perundang-undangan tanpa menggunakan *query expansion* dengan nilai *precision* sebesar 41,33%, nilai *recall* sebesar 41,33%, nilai *f-measure* sebesar 41,33% dan nilai *accuracy* sebesar 99,41%. Sedangkan pengembalian dokumen oleh sistem ketika menggunakan *query expansion* adalah nilai *precision* sebesar 43,67%, nilai *recall* sebesar 43,67%, nilai *f-measure* sebesar 43,67% dan nilai *accuracy* sebesar 99,44%. Peningkatan yang dihasilkan dari uji coba tanpa menggunakan *query expansion* dan uji coba menggunakan *query expansion* ditinjau dari parameter *precision*, *recall*, dan *f-measure* adalah sebesar 2,34%, serta *accuracy* sebesar 0,3%.

Error yang terdapat pada sistem ketika mengembalikan dokumen terjadi karena beberapa faktor, antara lain kekurangan pada algoritma proses *preprocessing* yang membuat *term* pada dokumen tidak diproses menjadi kata dasar dengan benar. Kesalahan ini menyebabkan *term* tidak terbaca sebagai *term* yang sama dengan *query* dan tidak terhitung sebagaimana mestinya sehingga mengurangi hasil dari *term probability*. *Term probability* yang tidak sesuai menyebabkan nilai *entropy* menjadi lebih rendah. Faktor kedua karena

kekurangan sistem dalam membaca dokumen PDF. Terdapat kekurangan pada sistem di mana sebuah kata pada PDF dibaca terpisah, seperti kata “mengurangi” menjadi “meng urangi”. Sedangkan dua kata atau lebih dibaca sebagai satu kata, seperti “menjadi sebuah negara” menjadi “menjadisebuahnegara”. Faktor ini menyebabkan perhitungan *term probability* menjadi lebih rendah karena *term* tidak terbaca dengan baik, sehingga nilai *entropy* yang dihasilkan juga semakin rendah.

Allah *subhanahu wa ta'ala* berfirman pada surah Al Furqan ayat 2 :

الَّذِي لَهُ مُلْكُ السَّمَوَاتِ وَالْأَرْضِ وَلَمْ يَتَّخِذْ وَلَدًا وَلَمْ يَكُنْ لَهُ شَرِيكٌ فِي الْمُلْكِ وَخَلَقَ كُلَّ شَيْءٍ فَقَدَرَهُ تَقْدِيرًا

“Yang memiliki kerajaan langit dan bumi, tidak mempunyai anak, tidak ada sekutu bagi-Nya dalam kekuasaan(-Nya), dan Dia menciptakan segala sesuatu, lalu menetapkan ukuran-ukurannya dengan tepat.”

Dalam tafsir Ibnu Katsir, ayat di atas menjelaskan bahwa Allah *subhanahu wa ta'ala* membersihkan diri-Nya dari beranak dan sekutu. Segala sesuatu adalah makhluk yang dimiliki oleh-Nya dan Dia adalah pencipta segala sesuatu, Yang Menguasai, Yang Memiliki dan Tuhannya, segala sesuatu berada di bawah kekuasaan-Nya, diatur oleh-Nya, tunduk kepada-Nya dan takdir-Nya.

Tujuan dari penelitian ini adalah untuk mengevaluasi kinerja sistem dalam mengembalikan dokumen yang relevan dengan *query* yang diberikan oleh pengguna menggunakan metode *Shannon's entropy*. Berdasarkan ayat di atas, Allah telah menyebutkan bahwa semua hal di dunia ini sudah memiliki ukurannya

masing-masing, begitu juga dengan dokumen yang terdapat pada repositori. Setiap dokumen akan memiliki nilai *entropy* masing-masing sesuai dengan *query* yang diberikan oleh pengguna.

Pada Al-Qur'an surah Al-Mujadalah ayat 11, Allah *subhanahu wa ta'ala* berfirman :

يَفْسَحْ فَأَفْسَحُوا الْمَجْلِسِ فِي تَفَسَّحُوا لَكُمْ قِيلَ إِذَا أَمُنُوا الَّذِينَ يَأْتِيهَا أَنْشُرُوا قِيلَ وَإِذَا لَكُمْ اللَّهُ
يَرْفَعْ فَأَنْشُرُوا خَبِيرٌ تَعْمَلُونَ بِمَا وَاللَّهُ دَرَجَاتٍ الْعِلْمِ أُوتُوا وَالَّذِينَ مِنْكُمْ أَمُنُوا الَّذِينَ اللَّهُ

“Wahai orang-orang yang beriman! Apabila dikatakan kepadamu, “Berilah kelapangan di dalam majelis-majelis,” maka lapangkanlah, niscaya Allah akan memberi kelapangan untukmu. Dan apabila dikatakan, “Berdirilah kamu,” maka berdirilah, niscaya Allah akan mengangkat (derajat) orang-orang yang beriman di antaramu dan orang-orang yang diberi ilmu beberapa derajat. Dan Allah Mahateliti apa yang kamu kerjakan”.

Dijelaskan dalam tafsir Ibnu Katsir, Allah *subhanahu wa ta'ala* memerintahkan kepada hamba-hambanya yang beriman supaya sebagian dari mereka bersikap baik terhadap sebagian yang lain pada sebuah majelis-majelis pertemuan, atau majelis yang mengajak kepada kebaikan. Perintah itu diberikan karena pembalasan sesuai dengan jenis amal perbuatan. Dan barangsiapa yang memberikan kelapangan atau kebaikan terhadap saudaranya ketika di dalam majelis, janganlah beranggapan bahwa itu akan mengurangi haknya dan merendharkannya. Tentunya tidak, melainkan hal itu dapat membuat derajatnya

tinggi di sisi Allah. Allah *subhanahu wa ta'ala* tidak akan menyia-nyiakan pahala untuknya dan akan memberikan pahalanya di dunia dan akhirat. Karena sesungguhnya barangsiapa yang berendah diri terhadap perintah Allah, niscaya Allah akan meninggikan kedudukannya dan mengharumkan namanya. Allah *subhanahu wa ta'ala* mengetahui siapa yang berhak untuk mendapatkannya dan siapa yang tidak berhak mendapatkannya.

Ayat di atas menjelaskan tentang orang yang beriman dan berilmu akan diangkat derajatnya oleh Allah. Semakin tinggi iman dan ilmu seseorang maka semakin tinggi pula derajat yang akan ia dapatkan. Hasil sistem yang dibangun sesuai dengan ayat di atas di mana terdapat perankingan dokumen sejumlah sepuluh teratas untuk mendapatkan hasil yang paling relevan. Semakin banyak *term* yang sama dengan *query* maka semakin relevan dokumen tersebut dan untuk dikembalikan juga semakin besar.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan uji coba yang telah dilakukan, kesimpulan yang didapatkan adalah sistem yang dibangun tanpa menggunakan *query expansion* memiliki nilai *precision* sebesar 41,33%, nilai *recall* sebesar 41,33%, nilai *f-measure* sebesar 41,33% dan nilai *accuracy* sebesar 99,41% dalam mengembalikan dokumen sejumlah 10 teratas terhadap masing-masing *query* yang berjumlah 30. Sedangkan sistem yang menggunakan *query expansion* menghasilkan nilai *precision* sebesar 43,67%, nilai *recall* sebesar 43,67%, nilai *f-measure* sebesar 43,67% dan nilai *accuracy* sebesar 99,44%. Nilai *precision*, *recall*, dan *f-measure* meningkat sebesar 2,34% dan nilai *accuracy* meningkat sebesar 0,3%. Proses *query expansion* menyebabkan *query* yang disusun kembali memiliki *term* tambahan yang berkaitan dengan *term* pada *query* awal sehingga pencarian dokumen menjadi lebih luas. Dokumen yang semula tidak dikategorikan sebagai dokumen relevan akan menjadi dokumen yang relevan karena memiliki *term* yang sama dengan *query* hasil *query expansion*. Sedangkan dokumen yang semula sudah dikategorikan sebagai dokumen relevan akan memiliki nilai relevanan yang lebih tinggi karena *term* yang sama dengan *query* menjadi bertambah. Hal inilah yang menyebabkan *performance* sistem dalam mengembalikan dokumen menggunakan metode *Shannon's entropy* dan *query expansion* menjadi lebih baik.

5.2 Saran

Dalam pengembangan sistem pengembalian dokumen perundang-undangan menggunakan *Shannon's entropy* dan *query expansion* ini memerlukan beberapa perbaikan untuk mencapai hasil yang lebih maksimal, maka saran yang mungkin dapat dilakukan untuk penelitian selanjutnya antara lain :

1. Menggunakan algoritma yang lebih baik untuk tahap *preprocessing* karena sangat mempengaruhi pembacaan *term* pada dokumen maupun *query*.
2. Membangun kamus yang lebih kaya istilah untuk mendapatkan hasil *query expansion* yang lebih baik. Seperti kamus antonim dan homonim.
3. Data yang digunakan bukan hanya dokumen perundang-undangan, namun dokumen pada bidang lain yang memerlukan pencarian menggunakan *query*.
4. Pendekatan metode lain untuk melakukan pencarian dokumen yang relevan dengan *query*.

DAFTAR PUSTAKA

- Acuna-ureta, D. E., Orchard, M. E., & Wheeler, P. (2020). Computation of time probability distributions for the occurrence of uncertain future events. *Mechanical Systems and Signal Processing* 150.
- Badan Pusat Statistik. (2019). *Statistik Kriminal 2019*. Jakarta : Badan Pusat Statistik.
- Balitbangkumham. (2016). *Kesadaran Hukum Masyarakat dalam Penerapan Undang-Undang Sistem Peradilan Pidana Anak*. Jakarta : Badan Penelitian dan Pengembangan Hukum dan HAM.
- Becker, M., Kasper, S., Bockmann, B., Jockel, K. H., & Virchow, I. (2019). Natural language processing of German clinical colorectal cancer notes for guideline-based treatment evaluation. *International Journal of Medical Informatics*, 127(February), 141-146.
- Bezadi, F. (2015). Natural language processing and machine learning : a review. *International Journal of Computer Science and Information Security*, 13(September), 101 – 106.
- Brownlee, J. (2019). A Gentle Introduction to Information Entropy. *Machine Learning Mastery* : October 14, 2019. <https://machinelearningmastery.com/what-is-information-entropy/>. Tanggal akses : 3 November 2020.
- Chen, C. B., & Wang, J. Y. (2006). Rough Set-Based Clustering with Refinement Using Shannon ' s Entropy Theory. *An International Journal Computers and Mathematics with Application*. 52, 1563-1576.
- CNNIndonesia. (2020). Kronologi 'Prank' Sembako ke Waria Youtuber Ferdian Paleka. Retrieved October 25, 2020, from : <https://www.cnnindonesia.com/>
- Fikri, A. D. (2019). “Perbandingan Metode Dice Similarity dengan Cosine Similarity Menggunakan Query Expansion pada Pencarian Ayatul Ahkam dalam Terjemah Alquran Berbahasa Indonesia”. Skripsi. Sains dan Teknologi. Teknik Informatika. Universitas Islam Negeri Maulana Malik Ibrahim. Malang.
- Gagie, T., Kärkkäinen, J., Navarro, G., & Puglisi, S. J. (2013). Colored range queries and document retrieval. *Theoretical Computer Science*, 483, 36–50.

- Iwanaga, J., Nishimura, N., Sukegawa, N., & Takano, Y. (2016). Estimating product-choice probabilities from recency and frequency of page views. *Knowledge-Based Systems*.
- Jatnika, D., Bijaksana, M. A., & Suryani, A. A. (2019). Word2vec model analysis for semantic similarities in English words. *Procedia Computer Science*, 157, 160–167.
- Kazemian-kale-kale, A., Bonakdari, H., & Gholami, A. (2019). The uncertainty of the Shannon entropy model for shear stress distribution in circular channels. *International Journal of Sediment Research*.
- Kivunja, C. (2018). *Distinguishing between Theory , Theoretical Framework , and Conceptual Framework : A Systematic Review of Lessons from the Field*. 7(6), 44–53.
- Liddy, E.D. (2001). Natural Language Processing. In *Encyclopedia of Library and Information Science, 2nd Ed.* NY. Marcel Decker, Inc.
- Ly, T., Pamer, C., Dang, O., Brajovic, S., Haider, S., Botsis, T., Milward, D., Winter, A., Lu, S., & Ball, R. (2018). Evaluation of Natural Language Processing (NLP) systems to annotate drug product labeling with MedDRA terminology. *Journal of Biomedical Informatics*, 83(December 2017), 73-86.
- Manning, C. D., Raghavan, P., Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge : Cambridge University Press.
- Mhamdi, F., & Elloumi, M. (2008). A new survey on knowledge discovery and data mining. *Proceedings of the 2nd International Conference on Research Challenges in Information Science, RCIS 2008, July*, 427-432.
- Mukaram, A. (2019). Ubah Pancasila Jadi Pancasila, Pemuda Kalbar Ditangkap Polisi. Retrieved October 25, 2020, from : <https://www.liputan6.com/>
- Nugraheny, D. E. (2019). Data Kependudukan 2020. Retrieved October 24, 2020, from : <https://www.kompas.com/>
- Ozgul, F. (2013). Chapter 9 - Incorporating Data and Methodologies for Knowledge Discovery for Crime. In *Intelligent Systems for Security Informatics*. Elsevier Inc.
- Ramaswamy, S., & DeClerck, N. (2018). Customer perception analysis using deep learning and NLP. *Procedia Computer Science*, 140, 170-178.

- Singh, J., & Sharan, A. (2015). Relevance feedback based query expansion model using borda count and semantic similarity approach. *Computational Intelligence and Neuroscience, 2015*.
- Tala, F.Z. (2003). *A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia*. Thesis. Institute for Logic Language and Computation Universiteit van Amsterdam The Netherlands.
- Walukow, J. M. (2013). Perwujudan Prinsip Equality Before the Law Bagi Narapidana di Dalam Lembaga Pemasyarakatan di Indonesia. *Lex et Societatis. 1 (1). 163 – 172*.
- Yousef, Nidal., Al-Bidewi, I. & Fayoumi, M.. (2010). Evaluation of different query expansion techniques and using different similarity measures in Arabic documents. *European Journal of Scientific Research, 43*. 156-166.
- Zhang, R., Guo, J., Fan, Y., Lan, Y., & Cheng, X. (2020). Query Understanding via Internet Description Generation. *Proceedings of the 29th ACM International Conference on Information and Knowledge Management, CIKM 2020, Oct*.

LAMPIRAN

Lampiran 1

Dokumen *groundtruth*

No	<i>Query</i>	Dokumen
1	Tunjangan hari raya untuk pegawai	PP No. 37 Tahun 2019
		PM Keuangan No. 49 Tahun 2020
		PP No. 24 Tahun 2020
		PM Keuangan No. 77 Tahun 2017
		PM Keuangan No. 54 Tahun 2018
		PM Keuangan No. 55 Tahun 2018
		PM Keuangan No. 56 Tahun 2018
		PP No. 22 Tahun 2016
		PP No. 25 Tahun 2017
		PP No. 36 Tahun 2019
2	Pemanfaatan hutan	PP No. 6 Tahun 2007
		PM Lingkungan Hidup No. 81 Tahun 2016
		PP No. 24 Tahun 2010
		PM Lingkungan Hidup No. 50 Tahun 2016
		UU No. 41 Tahun 1999
		PP No. 60 Tahun 2009
		PM Lingkungan Hidup No. 49 Tahun 2017
		PP No. 34 Tahun 2002
		PM Perhutanan No. 47 Tahun 2013
PP No. 3 Tahun 2008		
3	Penyelenggaraan penyiaran	PM Kominfo No. 22 Tahun 2011
		PP No. 52 Tahun 2005
		PP No. 51 Tahun 2005
		PP No. 50 Tahun 2005
		PM Kominfo No. 5 Tahun 2018
		PP No. 11 Tahun 2005
		PM Kominfo No. 32 Tahun 2007
		PM Kominfo No. 18 Tahun 2016
		PP No. 49 Tahun 2005
UU No. 32 Tahun 2002		
4	Jaringan telekomunikasi berbasis internet	PM Kominfo No. 16 Tahun 2010
		PM Kominfo No. 27 Tahun 2006
		PM Kominfo No. 5 Tahun 2017
		PM Kominfo No.8 Tahun 2017
		PM Kominfo No. 2 Tahun 2013
		PM Kominfo No. 1 Tahun 2013
		PM Kominfo No. 21 Tahun 2011
		PP No. 46 Tahun 2021
		PP No. 52 Tahun 2000
UU No. 11 Tahun 2008		

5	Persyaratan perizinan jasa penyiaran televisi	PM Kominfo No. 18 Tahun 2016
		PM Kominfo No. 28 Tahun 2013
		PM Kominfo No. 28 Tahun 2008
		PP No. 51 Tahun 2005
		PP No. 80 Tahun 2015
		PP No. 50 Tahun 2005
		PP No. 46 Tahun 2021
		PM Kominfo No. 41 Tahun 2012
		PP No. 24 Tahun 2018
		PP No. 7 Tahun 2009
		PM Kominfo No. 18 Tahun 2016
6	Ibadah Haji	UU No. 13 Tahun 2008
		PM Agama No. 20 Tahun 2016
		PM Agama No. 23 Tahun 2016
		UU No. 8 Tahun 2019
		UU No. 17 Tahun 1999
		PM Agama No. 7 Tahun 2019
		PP No. 79 Tahun 2012
		KM Agama 719 2020
		PM Agama No. 28 Tahun 2019
PM Agama No. 41 Tahun 2017		
7	Manfaat hutan untuk menanggulangi banjir dan erosi	PP No. 76 Tahun 2008
		PP No. 26 Tahun 2020
		PP No. 24 Tahun 2010
		UU No. 5 Tahun 1967
		PP No. 23 Tahun 2021
		PM Lingkungan Hidup No. 50 Tahun 2016
		UU No. 41 Tahun 1999
		PM Kehutanan No. 39 Tahun 2009
		PERPRES No. 28 Tahun 2011
		PM Lingkungan Hidup No. 32 Tahun 2015
8	Perkawinan	PP No. 9 Tahun 1975
		UU No. 1 Tahun 1974
		PP No. 10 Tahun 1983
		UU No. 16 Tahun 2019
		PP No. 45 Tahun 1990
		PM Agama No. 19 Tahun 2018
		PM Agama No. 11 Tahun 2007
		PM Agama No. 20 Tahun 2019
		UU No. 32 Tahun 1954
UU No. 22 Tahun 1946		
9	Fungsi pertahanan bagi sebuah negara	PERPRES No. 97 Tahun 2015
		PM Pertahanan No. 19 Tahun 2015
		PERPRES No. 7 Tahun 2008
		UU No. 20 Tahun 1982
		PM Pertahanan No. 27 Tahun 2013
		PM Pertahanan No. 16 Tahun 2019
PM Pertahanan No. 11 Tahun 2009		

		UU No. 3 Tahun 2002
		PM Pertahanan No. 82 Tahun 2014
		PERPRES No. 43 Tahun 2020
10	Gaji pegawai	PP No. 23 Tahun 1955
		PP No. 15 Tahun 2019
		PP No. 14 Tahun 1962
		PP No. 12 Tahun 1967
		PM Keuangan No. 138 Tahun 2011
		PERPRES No. 16 Tahun 2019
		PP No. 22 Tahun 2013
		PP No. 11 Tahun 2003
		PP No. 7 Tahun 1977
		PERPRES No. 1 Tahun 2006
11	Persyaratan penanaman modal usaha	PERPRES No. 44 Tahun 2016
		PERPRES No. 10 Tahun 2021
		PERPRES No. 39 Tahun 2014
		UU No. 25 Tahun 2007
		PERPRES No. 77 Tahun 2007
		PERPRES No. 76 Tahun 2007
		PERPRES No. 27 Tahun 2009
		PERPRES No. 16 Tahun 2012
		PP No. 45 Tahun 2008
		UU No. 11 Tahun 2020
12	Rehabilitasi wilayah pesisir	PERPRES No. 121 Tahun 2012
		PM Kelautan No. 24 Tahun 2016
		UU No. 27 Tahun 2007
		PM Kehutanan No. 35 Tahun 2010
		PM Kelautan No. 17 Tahun 2008
		UU No. 1 Tahun 2014
		PM Kelautan No. 4 Tahun 2018
		PM Kelautan No. 12 Tahun 2013
		PP No. 32 Tahun 2019
		PM Kehutanan No. 9 Tahun 2013
13	Misi pemeliharaan perdamaian	PM Pertahanan No. 14 Tahun 2016
		PM Pertahanan No. 11 Tahun 2013
		PERPRES No. 78 Tahun 2015
		PERPRES No. 86 Tahun 2015
		PERPRES No. 85 Tahun 2011
		PERPRES No. 50 Tahun 2010
		KEPPRES No. 7 Tahun 2019
		KEPPRES No. 4 Tahun 2008
		UU No. 37 Tahun 1999
		PERPRES No. 42 Tahun 2019
14	Penyelenggaraan perjalanan ibadah umrah	PM Agama No. 18 Tahun 2015
		PM Agama No 8 Tahun 2018
		KM Agama No. 719 Tahun 2020
		UU No. 8 Tahun 2019
		UU No. 17 Tahun 1999

		PP No. 79 Tahun 2012
		PP No. 38 Tahun 2021
		UU No. 13 Tahun 2008
		PM Agama No. 15 Tahun 2020
		PM Keuangan No. 92 Tahun 2020
15	Larangan mengeluarkan benih ikan dari Indonesia	PM Kelautan No. 21 Tahun 2014
		PM Kelautan No. 18 Tahun 2009
		PM Kelautan No. 50 Tahun 2017
		PM Kelautan No. 29 Tahun 2009
		UU No. 31 Tahun 2004
		PP No. 28 Tahun 2017
		PM Kelautan No. 18 Tahun 2020
		PM Kelautan No. 45 Tahun 2015
		UU No. 17 Tahun 2006
		UU No. 18 Tahun 2012
		PM Kelautan No. 21 Tahun 2014
16	Organisasi dan tata kerja balai penelitian	PM Kelautan No. 19 Tahun 2017
		PM Dikbud No. 27 Tahun 2015
		PM Kesehatan No. 65 Tahun 2017
		PM Lingkungan Hidup No. 24 Tahun 2016
		PM Kelautan No. 17 Tahun 2017
		PM Lingkungan Hidup No. 26 Tahun 2016
		PM Kelautan No. 34 Tahun 2011
		PM Kominfo No. 4 Tahun 2018
		PM Agama No. 18 Tahun 2018
		PM Dikbud No. 30 Tahun 2012
17	Ekspor dan impor narkoba	PM Kesehatan No. 10 Tahun 2013
		PM Kesehatan No. 3 Tahun 2015
		PP No. 40 Tahun 2013
		PP No. 44 Tahun 2010
		UU No. 35 Tahun 2009
		UU No. 10 Tahun 1995
		UU No. 17 Tahun 2006
		KM Kesehatan No. 825 Tahun 2008
		UU No. 22 Tahun 1997
		UU No. 9 Tahun 1976
18	Pengadaan vaksin Corona	PERPRES No. 99 Tahun 2020
		PM Kesehatan No. 10 Tahun 2021
		PERPRES No. 14 Tahun 2021
		PM Kesehatan No. 84 Tahun 2020
		PM Kesehatan No. 28 Tahun 2020
		UU No. 4 Tahun 1984
		UU No. 36 Tahun 2009
		KM Kesehatan No. 9860 Tahun 2020
		PM Keuangan No. 239 Tahun 2020
		PM Pertahanan No. 27 Tahun 2016
19	Pendirian rumah sakit umum	PM Kesehatan No. 30 Tahun 2019
		PM Kesehatan No. 3 Tahun 2020

		PM Kesehatan No. 56 Tahun 2014
		PP No. 47 Tahun 2021
		PM Kesehatan No. 24 Tahun 2016
		PM Kesehatan No. 340 Tahun 2010
		PM Kesehatan No. 147 Tahun 2010
		UU No. 44 Tahun 2009
		PM Kesehatan No. 11 Tahun 2016
		PM Kesehatan No. 1173 Tahun 2004
20	Bantuan kesehatan oleh pemerintah	PP No. 101 Tahun 2012
		PERPRES No. 82 Tahun 2018
		PM Kesehatan No. 28 Tahun 2014
		PM Kesehatan No. 307 Tahun 2009
		PERPRES No. 64 Tahun 2020
		PP No. 7 Tahun 2019
		PP No. 22 Tahun 2008
		PP No. 103 Tahun 2014
		UU No. 40 Tahun 2004
		PM Kesehatan No. 61 Tahun 2017
21	Budidaya perikanan	PP No. 28 Tahun 2017
		PM Kelautan No. 35 Tahun 2011
		PERPRES No. 2 Tahun 2017
		PM Kelautan No. 45 Tahun 2015
		PM Kelautan No. 7 Tahun 2018
		PM Kelautan No. 49 Tahun 2014
		PM Kelautan No. 12 Tahun 2011
		PERPRES No. 63 Tahun 2015
		PM Kelautan No. 15 Tahun 2005
		UU No. 45 Tahun 2009
22	Tarif layanan badan layanan umum	PM Keuangan No. 27 Tahun 2021
		PM Keuangan No. 100 Tahun 2016
		PM Keuangan No. 184 Tahun 2016
		PM Keuangan No. 264 Tahun 2016
		PP No. 74 Tahun 2012
		PM Keuangan No. 109 Tahun 2017
		PM Keuangan No. 216 Tahun 2011
		PM Keuangan No. 49 Tahun 2015
		PM Keuangan No. 142 Tahun 2020
		PM Keuangan No. 51 Tahun 2018
23	Anggaran pendapatan dan belanja negara	UU No. 9 Tahun 2020
		UU No. 20 Tahun 2019
		UU No. 45 Tahun 2007
		UU No. 9 Tahun 2005
		PERPRES No. 86 Tahun 2017
		PERPRES No. 129 Tahun 2018
		UU No. 15 Tahun 1968
		UU No. 14 Tahun 2015
		UU No. 18 Tahun 2016
		PP No. 45 Tahun 2013

24	Pengaturan penggunaan frekuensi radio	PM Kominfo No. 17 Tahun 2005
		PM Kominfo No. 5 Tahun 2019
		PP No. 53 Tahun 2000
		PM Kominfo No. 4 Tahun 2015
		PM Kominfo No. 3 Tahun 2016
		PM Kominfo No. 19 Tahun 2015
		PM Kominfo No. 1 Tahun 2006
		PM Kominfo No. 19 Tahun 2013
		UU No. 3 Tahun 1989
		PP No. 46 Tahun 2021
25	Pembiayaan telekomunikasi di Indonesia	PM Kominfo No. 10 Tahun 2013
		PM Kominfo No. 25 Tahun 2015
		PM Keuangan No. 85 Tahun 2018
		PM Kominfo No. 2 Tahun 2013
		PM Kominfo No. 10 Tahun 2018
		PM Kominfo No. 21 Tahun 2011
		PM Kominfo No. 3 Tahun 2018
		PM Keuangan No. 100 Tahun 2009
		PP No. 46 Tahun 2021
		PERPRES No. 32 Tahun 2020
26	Tarif sewa saluran siaran	PM Kominfo No. 18 Tahun 2012
		PM Kominfo No. 27 Tahun 2014
		PM Kominfo No. 28 Tahun 2013
		PM Kominfo No. 32 Tahun 2013
		PM Kominfo No. 39 Tahun 2009
		PM Kominfo No. 3 Tahun 2019
		PP No. 46 Tahun 2021
		PP No. 66 Tahun 2020
		PP No. 7 Tahun 2009
		PP No. 52 Tahun 2000
27	Universitas dan perguruan tinggi	UU No. 12 Tahun 2012
		PP No. 80 Tahun 2014
		PP No. 4 Tahun 2014
		UU No. 22 Tahun 1961
		PM Ristekdikti No. 51 Tahun 2018
		PP No. 60 Tahun 1999
		PP No. 46 Tahun 2019
		PM Ristekdikti No. 84 Tahun 2017
		PM Dikbud No. 25 Tahun 2020
		PP No. 30 Tahun 1990
28	Jaminan kesehatan	PERPRES No. 64 Tahun 2020
		PERPRES No. 25 Tahun 2020
		PERPRES No. 75 Tahun 2019
		UU No. 40 Tahun 2004
		UU No. 24 Tahun 2011
		PERPRES No. 82 Tahun 2018
PERPRES No. 12 Tahun 2013		

		PM Kesehatan No. 28 Tahun 2014
		PP No. 101 Tahun 2012
		PM Kesehatan No. 19 Tahun 2014
29	Organisasi dan tata kerja balai pengelolaan	PM Lingkungan Hidup No. 12 Tahun 2016
		PM Lingkungan Hidup No. 14 Tahun 2016
		PM Kehutanan No. 2 Tahun 2007
		PM Lingkungan Hidup No. 10 Tahun 2016
		PM Lingkungan Hidup No. 70 Tahun 2019
		PM Lingkungan Hidup No. 11 Tahun 2020
		PM Kehutanan No. 3 Tahun 2007
		PM Kehutanan No. 51 Tahun 2009
		PM Lingkungan Hidup No. 18 Tahun 2016
		PM Kehutanan No. 43 Tahun 2013
		30
PM Lingkungan Hidup No. 12 Tahun 2018		
PM Lingkungan Hidup No. 12 Tahun 2006		
PM Lingkungan Hidup No. 102 Tahun 2018		
PP No. 21 Tahun 2010		
PP No. 82 Tahun 2001		
PP No. 19 Tahun 1999		
PM Lingkungan Hidup No. 22 Tahun 2018		
PM Kelautan No. 24 Tahun 2019		
KM Lingkungan Hidup No. 259 Tahun 2018		

Lampiran 2

Hasil pengembalian dokumen tanpa menggunakan *query expansion*.

Query 1 : Tunjangan hari raya untuk pegawai.

No	Dokumen	Entropy
1	PP No. 37 Tahun 2019.pdf	0,6023
2	PM Keuangan No. 79 Tahun 2019.pdf	0,5698
3	PM Keuangan No. 56 Tahun 2018.pdf	0,5480
4	PM Keuangan No. 59 Tahun 2019.pdf	0,5440
5	PM Keuangan No. 58 Tahun 2019.pdf	0,5152
6	PP No. 22 Tahun 2016.pdf	0,4929
7	PM Keuangan No. 49 Tahun 2020.pdf	0,4893
8	PM Keuangan No. 54 Tahun 2018.pdf	0,4768
9	PP No. 20 Tahun 2018.pdf	0,4580
10	PM Kominfo No. 42 Tahun 2014.pdf	0,4575

Query 2 : Pemanfaatan hutan.

No	Dokumen	Entropy
1	PM Lingkungan Hidup No. 69 Tahun 2019.pdf	0,5188
2	PM Kehutanan No. 47 Tahun 2013.pdf	0,4947
3	PP No. 34 Tahun 2002.pdf	0,4745
4	PM Lingkungan Hidup No. 37 Tahun 2019.pdf	0,4584
5	PM Lingkungan Hidup No. 75 Tahun 2016.pdf	0,4538
6	PP No. 23 Tahun 2021.pdf	0,4494
7	PP No. 6 Tahun 2007.pdf	0,4489
8	UU No. 5 Tahun 1967.pdf	0,4314
9	PM Lingkungan Hidup No. 68 Tahun 2019.pdf	0,4308
10	UU No. 41 Tahun 1999.pdf	0,4260

Query 3 : Penyelenggaraan penyiaran.

No	Dokumen	Entropy
1	PM Kominfo No. 28 Tahun 2013.pdf	0,5477
2	PM Kominfo No. 40 Tahun 2012.pdf	0,5337
3	PM Kominfo No. 41 Tahun 2012.pdf	0,5010
4	PM Kominfo No. 22 Tahun 2011.pdf	0,4786
5	UU No. 24 Tahun 1997.pdf	0,4768
6	PM Kominfo No. 32 Tahun 2013.pdf	0,4761
7	PP No. 11 Tahun 2005.pdf	0,4478
8	PP No. 52 Tahun 2005.pdf	0,4454
9	PP No. 51 Tahun 2005.pdf	0,4445
10	PM Kominfo No. 39 Tahun 2012.pdf	0,4430

Query 4 : Jaringan telekomunikasi berbasis internet.

No	Dokumen	Entropy
1	PM Kominfo No. 27 Tahun 2006.pdf	0,6329
2	PM Kominfo No. 16 Tahun 2010.pdf	0,5485
3	PP No. 52 Tahun 2000.pdf	0,5370
4	PM Kominfo No. 29 Tahun 2010.pdf	0,5331
5	PM Kominfo No. 5 Tahun 2017.pdf	0,5314
6	PM Kominfo No. 8 Tahun 2017.pdf	0,4971
7	PM Kominfo No. 21 Tahun 2010.pdf	0,4857
8	UU No. 36 Tahun 1999.pdf	0,4567
9	PM Kominfo No. 24 Tahun 2011.pdf	0,4412
10	PM Kominfo No. 5 Tahun 2013.pdf	0,4389

Query 5 : Persyaratan perizinan jasa penyiaran televisi.

No	Dokumen	Entropy
1	PP No. 50 Tahun 2005.pdf	0,6027
2	PM Kominfo No. 28 Tahun 2013.pdf	0,5826
3	PP No. 52 Tahun 2005.pdf	0,5810
4	PP No. 51 Tahun 2005.pdf	0,5392
5	PP No. 11 Tahun 2005.pdf	0,5107
6	UU No. 24 Tahun 1997.pdf	0,5104
7	PM Kominfo No. 32 Tahun 2013.pdf	0,5038
8	PM Kominfo No. 22 Tahun 2011.pdf	0,5029
9	UU No. 32 Tahun 2002.pdf	0,4996
10	PM Kominfo No. 18 Tahun 2016.pdf	0,4992

Query 6 : Ibadah haji.

No	Dokumen	Entropy
1	PM Agama No. 13 Tahun 2016.pdf	0,5057
2	UU No. 17 Tahun 1999.pdf	0,4765
3	UU No. 13 Tahun 2008.pdf	0,4394
4	UU No. 34 Tahun 2009.pdf	0,4391
5	PM Agama No. 28 Tahun 2019.pdf	0,4266
6	PM Agama No. 20 Tahun 2016.pdf	0,4213
7	PM Agama No. 13 Tahun 2018.pdf	0,4113
8	PM Agama No. 11 Tahun 2018.pdf	0,3849
9	PP No. 79 Tahun 2012.pdf	0,3613
10	PM Agama No. 7 Tahun 2019.pdf	0,3608

Query 7 : Manfaat hutan untuk menanggulangi banjir dan erosi.

No	Dokumen	Entropy
1	PM Lingkungan Hidup No. 69 Tahun 2019.pdf	0,5188

2	PM Kehutanan No. 47 Tahun 2013.pdf	0,4947
3	PP No. 34 Tahun 2002.pdf	0,4772
4	PM Lingkungan Hidup No. 37 Tahun 2019.pdf	0,4584
5	UU No. 5 Tahun 1967.pdf	0,4570
6	PM Lingkungan Hidup No. 75 Tahun 2016.pdf	0,4538
7	PP No. 23 Tahun 2021.pdf	0,4525
8	PP No. 6 Tahun 2007.pdf	0,4489
9	UU No. 41 Tahun 1999.pdf	0,4395
10	PM Lingkungan Hidup No. 68 Tahun 2019.pdf	0,4308

Query 8 : Perkawinan.

No	Dokumen	Entropy
1	UU No. 1 Tahun 1974.pdf	0,2557
2	UU No. 16 Tahun 2019.pdf	0,1934
3	PP No. 9 Tahun 1975.pdf	0,1914
4	PM Agama No. 19 Tahun 2018.pdf	0,1848
5	PP No. 10 Tahun 1983.pdf	0,0928
6	PP No. 45 Tahun 1990.pdf	0,0807
7	UU No. 22 Tahun 1946.pdf	0,0636
8	UU No. 32 Tahun 1954.pdf	0,0488
9	PM Pertahanan No. 6 Tahun 2009.pdf	0,0467
10	UU No. 54 Tahun 1954.pdf	0,0436

Query 9 : Fungsi pertahanan bagi sebuah negara.

No	Dokumen	Entropy
1	PM Pertahanan No. 40 Tahun 2015.pdf	0,6997
2	PM Pertahanan No. 24 Tahun 2015.pdf	0,6755
3	PM Pertahanan No. 38 Tahun 2015.pdf	0,6686
4	PM Pertahanan No. 5 Tahun 2011.pdf	0,6581
5	PM Pertahanan No. 37 Tahun 2015.pdf	0,6528
6	PM Pertahanan No. 19 Tahun 2015.pdf	0,6421
7	PERPRES No. 41 Tahun 2010.pdf	0,6369
8	PM Pertahanan No. 1 Tahun 2010.pdf	0,6217
9	PM Pertahanan No. 20 Tahun 2011.pdf	0,6197
10	PERPRES No. 97 Tahun 2015.pdf	0,6064

Query 10 : Gaji pegawai.

No	Dokumen	Entropy
1	PERPRES No. 1 Tahun 2006.pdf	0,4825
2	PP No. 23 Tahun 1955.pdf	0,4596
3	PP No. 7 Tahun 1977.pdf	0,4570
4	PERPRES No. 16 Tahun 2019.pdf	0,4118
5	PP No. 12 Tahun 1967.pdf	0,3871

6	PP No. 14 Tahun 1962.pdf	0,3825
7	PP No. 200 Tahun 1961.pdf	0,3675
8	PP No. 11 Tahun 2003.pdf	0,3471
9	PP No. 34 Tahun 2014.pdf	0,3004
10	PP No. 15 Tahun 2019.pdf	0,2992

Query 11 : Persyaratan penanaman modal usaha.

No	Dokumen	<i>Entropy</i>
1	PERPRES No. 39 Tahun 2014.pdf	0,8849
2	PERPRES No. 77 Tahun 2007.pdf	0,8727
3	PERPRES No. 76 Tahun 2007.pdf	0,8566
4	UU No. 25 Tahun 2007.pdf	0,6782
5	PM Kelautan No. 30 Tahun 2009.pdf	0,6661
6	KEPPRES No. 34 Tahun 1992.pdf	0,6271
7	PERPRES No. 10 Tahun 2021.pdf	0,6046
8	KEPPRES No. 41 Tahun 1996.pdf	0,5475
9	PERPRES No. 16 Tahun 2012.pdf	0,5305
10	PERPRES No. 27 Tahun 2009.pdf	0,5288

Query 12 : Rehabilitasi wilayah pesisir.

No	Dokumen	<i>Entropy</i>
1	UU No. 27 Tahun 2007.pdf	0,3578
2	UU No. 1 Tahun 2014.pdf	0,3093
3	PERPRES No. 121 Tahun 2012.pdf	0,3060
4	PM Kelautan No. 8 Tahun 2018.pdf	0,2895
5	PM Kelautan No. 3 Tahun 2018.pdf	0,2858
6	PM Kelautan No. 40 Tahun 2014.pdf	0,2781
7	PM Pertahanan No. 13 Tahun 2014.pdf	0,2768
8	PM Kelautan No. 8 Tahun 2009.pdf	0,2705
9	PP No. 64 Tahun 2010.pdf	0,2697
10	KM Kelautan No. 33 Tahun 2002.pdf	0,2574

Query 13 : Misi pemeliharaan perdamaian.

No	Dokumen	<i>Entropy</i>
1	PM Pertahanan No. 14 Tahun 2016.pdf	0,4924
2	PM Pertahanan No. 30 Tahun 2011.pdf	0,4288
3	PERPRES No. 85 Tahun 2011.pdf	0,4283
4	PERPRES No. 50 Tahun 2010.pdf	0,4096
5	KEPPRES No. 7 Tahun 2019.pdf	0,3812
6	PERPRES No. 78 Tahun 2015.pdf	0,3023
7	PM Pertahanan No. 11 Tahun 2013.pdf	0,2991
8	PERPRES No. 128 Tahun 2014.pdf	0,2940
9	KEPPRES No. 4 Tahun 2008.pdf	0,2870

10	PM Pertahanan No. 21 Tahun 2009.pdf	0,2670
----	-------------------------------------	--------

Query 14 : Penyelenggaraan perjalanan ibadah umrah.

No	Dokumen	Entropy
1	PM Keuangan No. 92 Tahun 2020.pdf	0,6555
2	PM Agama No. 18 Tahun 2015.pdf	0,5526
3	UU No. 17 Tahun 1999.pdf	0,5413
4	PM Agama No. 8 Tahun 2018.pdf	0,4894
5	UU No. 13 Tahun 2008.pdf	0,4810
6	PM Agama No. 28 Tahun 2019.pdf	0,4375
7	UU No. 8 Tahun 2019.pdf	0,4264
8	PP No. 79 Tahun 2012.pdf	0,3910
9	PM Pertahanan No. 22 Tahun 2008.pdf	0,3848
10	UU No. 34 Tahun 2009.pdf	0,3661

Query 15 : Larangan mengeluarkan benih ikan dari Indonesia.

No	Dokumen	Entropy
1	PM Kelautan No. 18 Tahun 2009.pdf	0,8491
2	PM Kelautan No. 19 Tahun 2012.pdf	0,8105
3	PM Kelautan No. 21 Tahun 2014.pdf	0,6520
4	PM Kelautan No. 9 Tahun 2009.pdf	0,6439
5	PM Kelautan No. 11 Tahun 2007.pdf	0,6036
6	PM Kelautan No. 8 Tahun 2011.pdf	0,6002
7	PM Kelautan No. 18 Tahun 2020.pdf	0,5707
8	PM Kelautan No. 10 Tahun 2015.pdf	0,5518
9	PM Kelautan No. 5 Tahun 2012.pdf	0,5467
10	PM Kelautan No. 56 Tahun 2014.pdf	0,5409

Query 16 : Organisasi dan tata kerja balai penelitian.

No	Dokumen	Entropy
1	PM Lingkungan Hidup No. 25 Tahun 2016.pdf	0,7359
2	PM Lingkungan Hidup No. 22 Tahun 2016.pdf	0,6242
3	PM Lingkungan Hidup No. 26 Tahun 2016.pdf	0,6143
4	PM Dikbud No. 66 Tahun 2016.pdf	0,6125
5	PM Dikbud No. 35 Tahun 2016.pdf	0,6091
6	PM Kesehatan No. 65 Tahun 2017.pdf	0,5998
7	PM Lingkungan Hidup No. 24 Tahun 2016.pdf	0,5990
8	PM Lingkungan Hidup No. 23 Tahun 2016.pdf	0,5746
9	PM Lingkungan Hidup No. 21 Tahun 2016.pdf	0,5653
10	PM Lingkungan Hidup No. 20 Tahun 2016.pdf	0,5637

Query 17 : Ekspor dan impor narkotika.

No	Dokumen	Entropy
1	UU No. 35 Tahun 2009.pdf	0,3448
2	PM Keuangan No. 21 Tahun 2019.pdf	0,3339
3	UU No. 22 Tahun 1997.pdf	0,3170
4	PM Kesehatan No. 10 Tahun 2013.pdf	0,3115
5	PM Keuangan No. 26 Tahun 2020.pdf	0,3096
6	UU No. 9 Tahun 1976.pdf	0,2892
7	PP No. 29 Tahun 2017.pdf	0,2784
8	PM Dikbud No. 38 Tahun 2019.pdf	0,2771
9	PP No. 40 Tahun 2013.pdf	0,2645
10	PM Kesehatan No. 168 Tahun 2005.pdf	0,2346

Query 18 : Pengadaan vaksin corona.

No	Dokumen	Entropy
1	KM Kesehatan No. 9860 Tahun 2020.pdf	0,4348
2	PM Kesehatan No. 28 Tahun 2020.pdf	0,4317
3	PERPRES No. 99 Tahun 2020.pdf	0,3530
4	PERPRES No. 14 Tahun 2021.pdf	0,2745
5	PM Keuangan No. 5 Tahun 2019.pdf	0,2352
6	PM Kelautan No. 1 Tahun 2012.pdf	0,2320
7	PM Pertahanan No. 28 Tahun 2008.pdf	0,2319
8	PM Lingkungan Hidup No. 35 Tahun 2015.pdf	0,2138
9	PP No. 148 Tahun 2015.pdf	0,2125
10	PM Kesehatan No. 84 Tahun 2020.pdf	0,2015

Query 19 : Pendirian rumah sakit umum.

No	Dokumen	Entropy
1	PM Kesehatan No. 3 Tahun 2020.pdf	0,5306
2	UU No. 44 Tahun 2009.pdf	0,5220
3	UU No. 45 Tahun 1957.pdf	0,4996
4	PM Kesehatan No. 147 Tahun 2010.pdf	0,4983
5	UU No. 39 Tahun 1958.pdf	0,4951
6	PP No. 93 Tahun 2015.pdf	0,4857
7	UU No. 51 Tahun 1954.pdf	0,4724
8	KM Kesehatan No. 88 Tahun 2014.pdf	0,4643
9	KM Kesehatan No. 11 Tahun 2014.pdf	0,4464
10	PP No. 47 Tahun 2021.pdf	0,4382

Query 20 : Bantuan kesehatan oleh pemerintah.

No	Dokumen	Entropy
1	PM Kesehatan No. 20 Tahun 2020.pdf	0,5261
2	PP No. 7 Tahun 1987.pdf	0,5195
3	PM Kesehatan No. 60 Tahun 2013.pdf	0,5096

4	PP No. 101 Tahun 2012.pdf	0,5058
5	PP No. 49 Tahun 1952.pdf	0,4603
6	KM Kesehatan No. 247 Tahun 2013.pdf	0,4595
7	PM Ristekdikti No. 95 Tahun 2016.pdf	0,4572
8	KM Kesehatan No. 783 Tahun 2006.pdf	0,4513
9	PM Agama No. 21 Tahun 2019.pdf	0,4480
10	UU No. 9 Tahun 1960.pdf	0,4473

Query 21 : Budidaya perikanan.

No	Dokumen	<i>Entropy</i>
1	PM Kelautan No. 7 Tahun 2009.pdf	0,4241
2	PM Kelautan No. 15 Tahun 2005.pdf	0,4206
3	PM Kelautan No. 21 Tahun 2009.pdf	0,4051
4	PM Kelautan No. 3 Tahun 2007.pdf	0,4049
5	PM Kelautan No. 7 Tahun 2010.pdf	0,4002
6	PERPRES No. 63 Tahun 2015.pdf	0,3976
7	PM Kelautan No. 11 Tahun 2007.pdf	0,3841
8	PP No. 15 Tahun 1990.pdf	0,3764
9	PM Kelautan No. 12 Tahun 2011.pdf	0,3682
10	PM Kelautan No. 35 Tahun 2011.pdf	0,3664

Query 22 : Tarif layanan badan layanan umum.

No	Dokumen	<i>Entropy</i>
1	PM Keuangan No. 16 Tahun 2019.pdf	0,7576
2	PM Keuangan No. 39 Tahun 2019.pdf	0,6920
3	PM Keuangan No. 3 Tahun 2020.pdf	0,6893
4	PM Keuangan No. 2 Tahun 2020.pdf	0,6842
5	PM Keuangan No. 6 Tahun 2020.pdf	0,6840
6	PM Keuangan No. 29 Tahun 2019.pdf	0,6811
7	PM Keuangan No. 54 Tahun 2019.pdf	0,6783
8	PM Keuangan No. 53 Tahun 2019.pdf	0,6715
9	PM Keuangan No. 42 Tahun 2019.pdf	0,6467
10	PM Keuangan No. 10 Tahun 2019.pdf	0,6414

Query 23 : Anggaran pendapatan dan belanja negara.

No	Dokumen	<i>Entropy</i>
1	PP No. 82 Tahun 2013.pdf	0,7510
2	PM Keuangan No. 14 Tahun 2020.pdf	0,7477
3	PM Keuangan No. 47 Tahun 2019.pdf	0,7140
4	PERPRES No. 86 Tahun 2017.pdf	0,7137
5	UU No. 15 Tahun 1968.pdf	0,7119
6	PM Pertahanan No. 32 Tahun 2019.pdf	0,5850
7	UU No. 9 Tahun 2005.pdf	0,5840

8	PP No. 51 Tahun 2019.pdf	0,5441
9	PM Kominfo No. 9 Tahun 2012.pdf	0,5370
10	PM Keuangan No. 117 Tahun 2020.pdf	0,5282

Query 24 : Pengaturan penggunaan frekuensi radio.

No	Dokumen	Entropy
1	PM Kominfo No. 5 Tahun 2019.pdf	0,9044
2	PP No. 53 Tahun 2000.pdf	0,8893
3	PM Kominfo No. 22 Tahun 2014.pdf	0,8567
4	PM Kominfo No. 3 Tahun 2016.pdf	0,8475
5	PM Kominfo No. 4 Tahun 2020.pdf	0,8397
6	PM Kominfo No. 17 Tahun 2005.pdf	0,7966
7	PM Kominfo No. 32 Tahun 2012.pdf	0,7606
8	PM Kominfo No. 19 Tahun 2013.pdf	0,7535
9	PM Kominfo No. 28 Tahun 2014.pdf	0,7362
10	PM Kominfo No. 12 Tahun 2017.pdf	0,7223

Query 25 : Pembiayaan telekomunikasi di Indonesia.

No	Dokumen	Entropy
1	PP No. 52 Tahun 2000.pdf	0,4582
2	PM Kominfo No. 22 Tahun 2005.pdf	0,4366
3	PM Kominfo No. 2 Tahun 2020.pdf	0,4159
4	UU No. 36 Tahun 1999.pdf	0,4151
5	PM Kominfo No. 20 Tahun 2014.pdf	0,4119
6	PM Kominfo No. 3 Tahun 2020.pdf	0,4076
7	PM Dikbud No. 17 Tahun 2019.pdf	0,4034
8	PM Kominfo No. 19 Tahun 2012.pdf	0,4009
9	PM Kominfo No. 19 Tahun 2016.pdf	0,3991
10	PM Keuangan No. 85 Tahun 2018.pdf	0,3977

Query 26 : Tarif sewa saluran siaran.

No	Dokumen	Entropy
1	PM Kominfo No. 18 Tahun 2012.pdf	0,8544
2	PM Kominfo No. 27 Tahun 2014.pdf	0,6717
3	PM Kominfo No. 3 Tahun 2019.pdf	0,4949
4	PM Kominfo No. 28 Tahun 2013.pdf	0,4344
5	UU No. 24 Tahun 1997.pdf	0,4023
6	PM Kominfo No. 40 Tahun 2012.pdf	0,3999
7	PM Kominfo No. 22 Tahun 2011.pdf	0,3916
8	PM Kominfo No. 41 Tahun 2012.pdf	0,3871
9	PM Kominfo No. 32 Tahun 2013.pdf	0,3870
10	PP No. 52 Tahun 2005.pdf	0,3714

Query 27 : Universitas dan perguruan tinggi.

No	Dokumen	Entropy
1	PM Ristekdikti No. 76 Tahun 2017.pdf	0,5572
2	PP No. 80 Tahun 2014.pdf	0,5350
3	UU No. 22 Tahun 1961.pdf	0,5344
4	PM Ristekdikti No. 62 Tahun 2016.pdf	0,4831
5	PM Ristekdikti No. 59 Tahun 2017.pdf	0,4816
6	PM Ristekdikti No. 11 Tahun 2018.pdf	0,4743
7	PM Ristekdikti No. 52 Tahun 2018.pdf	0,4675
8	KM Ristekdikti No. 205 Tahun 2018.pdf	0,4637
9	PM Ristekdikti No. 91 Tahun 2017.pdf	0,4489
10	PM Ristekdikti No. 6 Tahun 2016.pdf	0,4387

Query 28 : Jaminan kesehatan.

No	Dokumen	Entropy
1	PM Kesehatan No. 13 Tahun 2019.pdf	0,5218
2	PP No. 101 Tahun 2012.pdf	0,5034
3	KM Kesehatan No. 455 Tahun 2013.pdf	0,4558
4	PERPRES No. 12 Tahun 2013.pdf	0,4513
5	PP No. 84 Tahun 2015.pdf	0,4469
6	PM Kesehatan No. 28 Tahun 2014.pdf	0,4131
7	PM Kesehatan No. 71 Tahun 2013.pdf	0,4023
8	PM Kesehatan No. 19 Tahun 2014.pdf	0,3979
9	PP No. 82 Tahun 2013.pdf	0,3976
10	PERPRES No. 10 Tahun 2009.pdf	0,3927

Query 29 : Organisasi dan tata kerja balai pengelolaan hutan.

No	Dokumen	Entropy
1	PM Lingkungan Hidup No. 12 Tahun 2016.pdf	0,8867
2	PM Lingkungan Hidup No. 25 Tahun 2016.pdf	0,8141
3	PM Lingkungan Hidup No. 10 Tahun 2016.pdf	0,7841
4	PM Lingkungan Hidup No. 6 Tahun 2016.pdf	0,7527
5	PM Kehutanan No. 43 Tahun 2013.pdf	0,7505
6	PM Lingkungan Hidup No. 14 Tahun 2016.pdf	0,7077
7	PM Lingkungan Hidup No. 16 Tahun 2016.pdf	0,6899
8	PM Lingkungan Hidup No. 23 Tahun 2016.pdf	0,6724
9	PM Lingkungan Hidup No. 26 Tahun 2016.pdf	0,6719
10	PM Lingkungan Hidup No. 24 Tahun 2016.pdf	0,6567

Query 30 : Tata cara pembuangan limbah ke laut.

No	Dokumen	Entropy
1	PM Lingkungan Hidup No. 12 Tahun 2006.pdf	0,5680

2	PM Lingkungan Hidup No. 12 Tahun 2018.pdf	0,5632
3	PM Lingkungan Hidup No. 102 Tahun 2018.pdf	0,4879
4	KM Lingkungan Hidup No. 259 Tahun 2018.pdf	0,4732
5	PM Lingkungan Hidup No. 20 Tahun 2020.pdf	0,4512
6	PP No. 19 Tahun 1999.pdf	0,4380
7	PM Lingkungan Hidup No. 54 Tahun 2017.pdf	0,4000
8	PM Lingkungan Hidup No. 5 Tahun 2010.pdf	0,3483
9	PM Lingkungan Hidup No. 6 Tahun 2010.pdf	0,3404
10	PM Lingkungan Hidup No. 68 Tahun 2016.pdf	0,3295

Lampiran 3

Hasil pengembalian dokumen menggunakan *query expansion*.

Query 1 : Tunjangan hari raya untuk pegawai.

No	Dokumen	Entropy
1	PP No. 37 Tahun 2019.pdf	0,6023
2	PM Keuangan No. 79 Tahun 2019.pdf	0,5698
3	PM Keuangan No. 56 Tahun 2018.pdf	0,5480
4	PM Keuangan No. 59 Tahun 2019.pdf	0,5440
5	PM Keuangan No. 58 Tahun 2019.pdf	0,5152
6	PP No. 22 Tahun 2016.pdf	0,4929
7	PM Keuangan No. 49 Tahun 2020.pdf	0,4893
8	PM Keuangan No. 54 Tahun 2018.pdf	0,4768
9	PP No. 20 Tahun 2018.pdf	0,4580
10	PM Kominfo No. 42 Tahun 2014.pdf	0,4575

Query 2 : Pemanfaatan hutan.

No	Dokumen	Entropy
1	PM Lingkungan Hidup No. 37 Tahun 2019.pdf	0,6037
2	PM Lingkungan Hidup No. 69 Tahun 2019.pdf	0,5362
3	PM Kehutanan No. 47 Tahun 2013.pdf	0,5265
4	PP No. 34 Tahun 2002.pdf	0,5127
5	PP No. 23 Tahun 2021.pdf	0,5017
6	UU No. 41 Tahun 1999.pdf	0,4902
7	PP No. 10 Tahun 2010.pdf	0,4862
8	PM Lingkungan Hidup No. 45 Tahun 2016.pdf	0,4837
9	PP No. 6 Tahun 2007.pdf	0,4826
10	UU No. 5 Tahun 1967.pdf	0,4779

Query 3 : Penyelenggaraan penyiaran.

No	Dokumen	Entropy
1	PM Kominfo No. 28 Tahun 2013.pdf	0,5917
2	PM Kominfo No. 22 Tahun 2011.pdf	0,5651
3	PM Kominfo No. 40 Tahun 2012.pdf	0,5642
4	UU No. 24 Tahun 1997.pdf	0,5528
5	PM Kominfo No. 32 Tahun 2013.pdf	0,5462
6	PM Kominfo No. 41 Tahun 2012.pdf	0,5293
7	PP No. 11 Tahun 2005.pdf	0,5062
8	PP No. 51 Tahun 2005.pdf	0,4950
9	PP No. 52 Tahun 2005.pdf	0,4913
10	PM Kominfo No. 39 Tahun 2012.pdf	0,4907

Query 4 : Jaringan telekomunikasi berbasis internet.

No	Dokumen	Entropy
1	PM Kominfo No. 27 Tahun 2006.pdf	0,6329
2	PM Kominfo No. 16 Tahun 2010.pdf	0,5485
3	PP No. 52 Tahun 2000.pdf	0,5370
4	PM Kominfo No. 29 Tahun 2010.pdf	0,5331
5	PM Kominfo No. 5 Tahun 2017.pdf	0,5314
6	PM Kominfo No. 8 Tahun 2017.pdf	0,4971
7	PM Kominfo No. 21 Tahun 2010.pdf	0,4857
8	UU No. 36 Tahun 1999.pdf	0,4567
9	PM Kominfo No. 24 Tahun 2011.pdf	0,4412
10	PM Kominfo No. 5 Tahun 2013.pdf	0,4389

Query 5 : Persyaratan perizinan jasa penyiaran televisi.

No	Dokumen	Entropy
1	PP No. 50 Tahun 2005.pdf	0,6158
2	PM Kominfo No. 28 Tahun 2013.pdf	0,6009
3	PP No. 52 Tahun 2005.pdf	0,5968
4	PP No. 51 Tahun 2005.pdf	0,5524
5	UU No. 24 Tahun 1997.pdf	0,5448
6	PP No. 11 Tahun 2005.pdf	0,5300
7	PM Kominfo No. 32 Tahun 2013.pdf	0,5190
8	PM Kominfo No. 41 Tahun 2012.pdf	0,5186
9	PM Kominfo No. 22 Tahun 2011.pdf	0,5138
10	PM Kominfo No. 18 Tahun 2016.pdf	0,5132

Query 6 : Ibadah haji.

No	Dokumen	Entropy
1	PM Agama No. 13 Tahun 2016.pdf	0,5057
2	UU No. 17 Tahun 1999.pdf	0,4765
3	UU No. 13 Tahun 2008.pdf	0,4394
4	UU No. 34 Tahun 2009.pdf	0,4391
5	PM Agama No. 28 Tahun 2019.pdf	0,4266
6	PM Agama No. 20 Tahun 2016.pdf	0,4213
7	PM Agama No. 13 Tahun 2018.pdf	0,4113
8	PM Agama No. 11 Tahun 2018.pdf	0,3849
9	PP No. 79 Tahun 2012.pdf	0,3613
10	PM Agama No. 7 Tahun 2019.pdf	0,3608

Query 7 : Manfaat hutan untuk menanggulangi banjir dan erosi.

No	Dokumen	Entropy
1	PM Lingkungan Hidup No. 37 Tahun 2019.pdf	0,6037

2	PM Lingkungan Hidup No. 69 Tahun 2019.pdf	0,5362
3	PM Kehutanan No. 47 Tahun 2013.pdf	0,5265
4	PP No. 34 Tahun 2002.pdf	0,5153
5	PP No. 23 Tahun 2021.pdf	0,5048
6	UU No. 41 Tahun 1999.pdf	0,5036
7	UU No. 5 Tahun 1967.pdf	0,5036
8	PP No. 10 Tahun 2010.pdf	0,4925
9	PM Lingkungan Hidup No. 45 Tahun 2016.pdf	0,4837
10	PP No. 6 Tahun 2007.pdf	0,4826

Query 8 : Perkawinan.

No	Dokumen	Entropy
1	UU No. 1 Tahun 1974.pdf	0,2557
2	UU No. 16 Tahun 2019.pdf	0,1934
3	PP No. 9 Tahun 1975.pdf	0,1914
4	PM Agama No. 19 Tahun 2018.pdf	0,1848
5	PP No. 10 Tahun 1983.pdf	0,0928
6	PP No. 45 Tahun 1990.pdf	0,0807
7	UU No. 22 Tahun 1946.pdf	0,0636
8	UU No. 32 Tahun 1954.pdf	0,0488
9	PM Pertahanan No. 6 Tahun 2009.pdf	0,0467
10	UU No. 54 Tahun 1954.pdf	0,0436

Query 9 : Fungsi pertahanan bagi sebuah negara.

No	Dokumen	Entropy
1	PM Pertahanan No. 40 Tahun 2015.pdf	0,7372
2	PM Pertahanan No. 24 Tahun 2015.pdf	0,6755
3	PM Pertahanan No. 38 Tahun 2015.pdf	0,6686
4	PM Pertahanan No. 5 Tahun 2011.pdf	0,6581
5	PM Pertahanan No. 37 Tahun 2015.pdf	0,6528
6	PM Pertahanan No. 19 Tahun 2015.pdf	0,6421
7	PERPRES No. 97 Tahun 2015.pdf	0,6381
8	PERPRES No. 41 Tahun 2010.pdf	0,6369
9	PM Pertahanan No. 1 Tahun 2010.pdf	0,6217
10	PM Pertahanan No. 20 Tahun 2011.pdf	0,6197

Query 10 : Gaji pegawai.

No	Dokumen	Entropy
1	PP No. 23 Tahun 1955.pdf	0,4885
2	PP No. 7 Tahun 1977.pdf	0,4877
3	PERPRES No. 1 Tahun 2006.pdf	0,4825
4	PP No. 14 Tahun 1962.pdf	0,4784
5	PP No. 41 Tahun 2016.pdf	0,4630

6	PP No. 12 Tahun 1967.pdf	0,4584
7	PP No. 21 Tahun 2016.pdf	0,4397
8	PERPRES No. 16 Tahun 2019.pdf	0,4118
9	PP No. 200 Tahun 1961.pdf	0,4020
10	PP No. 11 Tahun 2003.pdf	0,3889

Query 11 : Persyaratan penanaman modal usaha.

No	Dokumen	<i>Entropy</i>
1	PERPRES No. 76 Tahun 2007.pdf	0,8903
2	PERPRES No. 39 Tahun 2014.pdf	0,8849
3	PERPRES No. 77 Tahun 2007.pdf	0,8727
4	UU No. 25 Tahun 2007.pdf	0,7563
5	PM Kelautan No. 30 Tahun 2009.pdf	0,6661
6	KEPPRES No. 34 Tahun 1992.pdf	0,6271
7	PERPRES No. 10 Tahun 2021.pdf	0,6194
8	KEPPRES No. 41 Tahun 1996.pdf	0,6056
9	PERPRES No. 27 Tahun 2009.pdf	0,5390
10	PP No. 45 Tahun 2008.pdf	0,5381

Query 12 : Rehabilitasi wilayah pesisir.

No	Dokumen	<i>Entropy</i>
1	UU No. 27 Tahun 2007.pdf	0,5532
2	PM Kelautan No. 8 Tahun 2009.pdf	0,4998
3	KM Kelautan No. 33 Tahun 2002.pdf	0,4782
4	UU No. 1 Tahun 2014.pdf	0,4667
5	PERPRES No. 121 Tahun 2012.pdf	0,4661
6	KEPPRES No. 114 Tahun 1999.pdf	0,4448
7	PM Kelautan No. 40 Tahun 2014.pdf	0,4403
8	PM Kelautan No. 3 Tahun 2018.pdf	0,4390
9	KM Agraria No. 6 Tahun 1990.pdf	0,4355
10	PM Kelautan No. 17 Tahun 2008.pdf	0,4307

Query 13 : Misi pemeliharaan perdamaian.

No	Dokumen	<i>Entropy</i>
1	PM Pertahanan No. 14 Tahun 2016.pdf	0,5769
2	PERPRES No. 50 Tahun 2010.pdf	0,5547
3	KEPPRES No. 7 Tahun 2019.pdf	0,5128
4	PERPRES No. 85 Tahun 2011.pdf	0,5124
5	KEPPRES No. 4 Tahun 2008.pdf	0,4980
6	PM Pertahanan No. 11 Tahun 2013.pdf	0,4365
7	PM Pertahanan No. 30 Tahun 2011.pdf	0,4288
8	PERPRES No. 78 Tahun 2015.pdf	0,4062
9	PERPRES No. 128 Tahun 2014.pdf	0,3923

10	PM Pertahanan No. 15 Tahun 2011.pdf	0,3433
----	-------------------------------------	--------

Query 14 : Penyelenggaraan perjalanan ibadah umrah.

No	Dokumen	Entropy
1	PM Keuangan No. 92 Tahun 2020.pdf	0,7058
2	UU No. 17 Tahun 1999.pdf	0,6320
3	PM Agama No. 18 Tahun 2015.pdf	0,5976
4	UU No. 13 Tahun 2008.pdf	0,5848
5	PM Pertahanan No. 22 Tahun 2008.pdf	0,5688
6	PM Agama No. 8 Tahun 2018.pdf	0,5385
7	PM Agama No. 28 Tahun 2019.pdf	0,5207
8	PP No. 79 Tahun 2012.pdf	0,5199
9	UU No. 8 Tahun 2019.pdf	0,5106
10	PM Pertahanan No. 28 Tahun 2008.pdf	0,4598

Query 15 : Larangan mengeluarkan benih ikan dari Indonesia.

No	Dokumen	Entropy
1	PM Kelautan No. 18 Tahun 2009.pdf	0,7198
2	PM Kelautan No. 19 Tahun 2012.pdf	0,7052
3	PERPRES No. 32 Tahun 2007.pdf	0,6179
4	PM Kelautan No. 8 Tahun 2011.pdf	0,6002
5	PM Kelautan No. 11 Tahun 2007.pdf	0,5959
6	PM Kelautan No. 21 Tahun 2014.pdf	0,5848
7	PM Kelautan No. 9 Tahun 2009.pdf	0,5531
8	PM Kelautan No. 10 Tahun 2015.pdf	0,5518
9	PM Kelautan No. 5 Tahun 2012.pdf	0,5467
10	PM Kelautan No. 56 Tahun 2014.pdf	0,5409

Query 16 : Organisasi dan tata kerja balai penelitian.

No	Dokumen	Entropy
1	PM Lingkungan Hidup No. 25 Tahun 2016.pdf	0,7519
2	PM Dikbud No. 66 Tahun 2016.pdf	0,6745
3	PM Kesehatan No. 65 Tahun 2017.pdf	0,6662
4	PM Lingkungan Hidup No. 22 Tahun 2016.pdf	0,6491
5	PM Lingkungan Hidup No. 26 Tahun 2016.pdf	0,6379
6	PM Lingkungan Hidup No. 24 Tahun 2016.pdf	0,6232
7	PM Dikbud No. 35 Tahun 2016.pdf	0,6091
8	PM Lingkungan Hidup No. 21 Tahun 2016.pdf	0,5909
9	PM Lingkungan Hidup No. 23 Tahun 2016.pdf	0,5908
10	PM Lingkungan Hidup No. 20 Tahun 2016.pdf	0,5882

Query 17 : Ekspor dan impor narkoba.

No	Dokumen	Entropy
1	PM Keuangan No. 26 Tahun 2020.pdf	0,3562
2	UU No. 35 Tahun 2009.pdf	0,3527
3	PM Keuangan No. 21 Tahun 2019.pdf	0,3466
4	PM Keuangan No. 36 Tahun 2019.pdf	0,3386
5	UU No. 22 Tahun 1997.pdf	0,3283
6	PM Keuangan No. 17 Tahun 2020.pdf	0,3243
7	PM Keuangan No. 31 Tahun 2020.pdf	0,3164
8	PM Keuangan No. 24 Tahun 2019.pdf	0,3161
9	PM Kesehatan No. 10 Tahun 2013.pdf	0,3155
10	UU No. 9 Tahun 1976.pdf	0,3017

Query 18 : Pengadaan vaksin corona.

No	Dokumen	Entropy
1	KM Kesehatan No. 9860 Tahun 2020.pdf	0,4348
2	PM Kesehatan No. 28 Tahun 2020.pdf	0,4317
3	PERPRES No. 99 Tahun 2020.pdf	0,3530
4	PERPRES No. 14 Tahun 2021.pdf	0,2745
5	PM Keuangan No. 5 Tahun 2019.pdf	0,2352
6	PM Kelautan No. 1 Tahun 2012.pdf	0,2320
7	PM Pertahanan No. 28 Tahun 2008.pdf	0,2319
8	PM Lingkungan Hidup No. 35 Tahun 2015.pdf	0,2138
9	PP No. 148 Tahun 2015.pdf	0,2125
10	PM Kesehatan No. 84 Tahun 2020.pdf	0,2015

Query 19 : Pendirian rumah sakit umum.

No	Dokumen	Entropy
1	UU No. 44 Tahun 2009.pdf	0,5014
2	UU No. 45 Tahun 1957.pdf	0,4996
3	UU No. 39 Tahun 1958.pdf	0,4951
4	PP No. 93 Tahun 2015.pdf	0,4857
5	PM Kesehatan No. 3 Tahun 2020.pdf	0,4729
6	UU No. 51 Tahun 1954.pdf	0,4724
7	KM Kesehatan No. 88 Tahun 2014.pdf	0,4643
8	KM Kesehatan No. 11 Tahun 2014.pdf	0,4464
9	PP No. 47 Tahun 2021.pdf	0,4340
10	PM Kesehatan No. 147 Tahun 2010.pdf	0,4183

Query 20 : Bantuan kesehatan oleh pemerintah.

No	Dokumen	Entropy
1	PP No. 7 Tahun 1987.pdf	0,5384
2	PM Kesehatan No. 20 Tahun 2020.pdf	0,5261
3	PM Kesehatan No. 60 Tahun 2013.pdf	0,5096

4	PP No. 101 Tahun 2012.pdf	0,5058
5	PP No. 49 Tahun 1952.pdf	0,4672
6	UU No. 36 Tahun 2009.pdf	0,4599
7	KM Kesehatan No. 247 Tahun 2013.pdf	0,4595
8	PM Ristekdikti No. 95 Tahun 2016.pdf	0,4572
9	UU No. 23 Tahun 1992.pdf	0,4570
10	UU No. 9 Tahun 1960.pdf	0,4558

Query 21 : Budidaya perikanan.

No	Dokumen	<i>Entropy</i>
1	PM Kelautan No. 35 Tahun 2011.pdf	0,4784
2	PERPRES No. 63 Tahun 2015.pdf	0,4604
3	PM Kelautan No. 7 Tahun 2009.pdf	0,4414
4	PM Kelautan No. 15 Tahun 2005.pdf	0,4330
5	PM Kelautan No. 21 Tahun 2009.pdf	0,4157
6	PERPRES No. 2 Tahun 2017.pdf	0,4148
7	PM Kelautan No. 31 Tahun 2011.pdf	0,4120
8	PM Kelautan No. 11 Tahun 2012.pdf	0,4075
9	PM Kelautan No. 11 Tahun 2007.pdf	0,4057
10	PM Kelautan No. 3 Tahun 2007.pdf	0,4049

Query 22 : Tarif layanan badan layanan umum.

No	Dokumen	<i>Entropy</i>
1	PM Keuangan No. 16 Tahun 2019.pdf	0,8255
2	PM Keuangan No. 85 Tahun 2018.pdf	0,7799
3	PM Keuangan No. 2 Tahun 2020.pdf	0,7772
4	PM Keuangan No. 3 Tahun 2020.pdf	0,7599
5	PM Keuangan No. 39 Tahun 2019.pdf	0,7551
6	PM Keuangan No. 54 Tahun 2019.pdf	0,7438
7	PM Keuangan No. 100 Tahun 2016.pdf	0,7302
8	PM Keuangan No. 29 Tahun 2019.pdf	0,7269
9	PM Keuangan No. 42 Tahun 2019.pdf	0,7184
10	PM Keuangan No. 10 Tahun 2019.pdf	0,7076

Query 23 : Anggaran pendapatan dan belanja negara.

No	Dokumen	<i>Entropy</i>
1	PM Keuangan No. 14 Tahun 2020.pdf	0,7780
2	PM Keuangan No. 47 Tahun 2019.pdf	0,7510
3	PERPRES No. 86 Tahun 2017.pdf	0,6953
4	PP No. 82 Tahun 2013.pdf	0,6125
5	PM Pertahanan No. 32 Tahun 2019.pdf	0,5850
6	UU No. 15 Tahun 1968.pdf	0,5724
7	PM Keuangan No. 117 Tahun 2020.pdf	0,5722

8	PM Keuangan No. 138 Tahun 2011.pdf	0,5345
9	PM Keuangan No. 135 Tahun 2020.pdf	0,5260
10	PM Kominfo No. 9 Tahun 2012.pdf	0,5145

Query 24 : Pengaturan penggunaan frekuensi radio.

No	Dokumen	Entropy
1	PM Kominfo No. 5 Tahun 2019.pdf	0,7031
2	PM Kominfo No. 4 Tahun 2020.pdf	0,7031
3	PM Kominfo No. 3 Tahun 2016.pdf	0,7001
4	PM Kominfo No. 19 Tahun 2013.pdf	0,6940
5	PM Kominfo No. 22 Tahun 2014.pdf	0,6921
6	PP No. 53 Tahun 2000.pdf	0,6807
7	PM Kominfo No. 17 Tahun 2005.pdf	0,6702
8	PM Kominfo No. 28 Tahun 2014.pdf	0,6494
9	PM Kominfo No. 32 Tahun 2012.pdf	0,6428
10	PM Kominfo No. 12 Tahun 2017.pdf	0,6243

Query 25 : Pembiayaan telekomunikasi di Indonesia.

No	Dokumen	Entropy
1	PM Keuangan No. 85 Tahun 2018.pdf	0,5921
2	PERPRES No. 86 Tahun 2017.pdf	0,5749
3	PP No. 22 Tahun 2015.pdf	0,5368
4	PM Keuangan No. 63 Tahun 2020.pdf	0,5158
5	PP No. 52 Tahun 2000.pdf	0,5073
6	PM Keuangan No. 53 Tahun 2019.pdf	0,5046
7	PM Kominfo No. 19 Tahun 2016.pdf	0,5031
8	PM Keuangan No. 14 Tahun 2020.pdf	0,4966
9	PERPRES No. 24 Tahun 2008.pdf	0,4952
10	PM Kominfo No. 22 Tahun 2005.pdf	0,4947

Query 26 : Tarif sewa saluran siaran.

No	Dokumen	Entropy
1	PM Kominfo No. 18 Tahun 2012.pdf	0,9425
2	PM Kominfo No. 27 Tahun 2014.pdf	0,7008
3	PM Kominfo No. 3 Tahun 2019.pdf	0,5219
4	PM Kominfo No. 28 Tahun 2013.pdf	0,4852
5	UU No. 24 Tahun 1997.pdf	0,4563
6	PM Kominfo No. 40 Tahun 2012.pdf	0,4365
7	PM Kominfo No. 41 Tahun 2012.pdf	0,4228
8	PP No. 11 Tahun 2005.pdf	0,4212
9	PP No. 52 Tahun 2005.pdf	0,4209
10	PM Kominfo No. 22 Tahun 2011.pdf	0,4079

Query 27 : Universitas dan perguruan tinggi.

No	Dokumen	Entropy
1	UU No. 22 Tahun 1961.pdf	0,5919
2	PM Ristekdikti No. 76 Tahun 2017.pdf	0,5779
3	PP No. 80 Tahun 2014.pdf	0,5350
4	PP No. 30 Tahun 1990.pdf	0,5103
5	PP No. 60 Tahun 1999.pdf	0,5080
6	PM Ristekdikti No. 52 Tahun 2018.pdf	0,4993
7	KM Ristekdikti No. 205 Tahun 2018.pdf	0,4854
8	PM Ristekdikti No. 62 Tahun 2016.pdf	0,4831
9	PM Ristekdikti No. 59 Tahun 2017.pdf	0,4816
10	PM Ristekdikti No. 11 Tahun 2018.pdf	0,4743

Query 28 : Jaminan kesehatan.

No	Dokumen	Entropy
1	PM Kesehatan No. 13 Tahun 2019.pdf	0,5218
2	PP No. 101 Tahun 2012.pdf	0,5034
3	PERPRES No. 12 Tahun 2013.pdf	0,4697
4	KM Kesehatan No. 455 Tahun 2013.pdf	0,4558
5	PP No. 84 Tahun 2015.pdf	0,4469
6	PM Kesehatan No. 28 Tahun 2014.pdf	0,4275
7	PM Kesehatan No. 71 Tahun 2013.pdf	0,4215
8	PERPRES No. 10 Tahun 2009.pdf	0,4148
9	PERPRES No. 82 Tahun 2018.pdf	0,4107
10	PM Kesehatan No. 19 Tahun 2014.pdf	0,3979

Query 29 : Organisasi dan tata kerja balai pengelolaan hutan.

No	Dokumen	Entropy
1	PM Lingkungan Hidup No. 12 Tahun 2016.pdf	0,8947
2	PM Lingkungan Hidup No. 25 Tahun 2016.pdf	0,8301
3	PM Lingkungan Hidup No. 10 Tahun 2016.pdf	0,8105
4	PM Kehutanan No. 43 Tahun 2013.pdf	0,7661
5	PM Lingkungan Hidup No. 6 Tahun 2016.pdf	0,7606
6	PM Lingkungan Hidup No. 14 Tahun 2016.pdf	0,7165
7	PM Lingkungan Hidup No. 16 Tahun 2016.pdf	0,7142
8	PM Lingkungan Hidup No. 26 Tahun 2016.pdf	0,6955
9	PM Keuangan No. 53 Tahun 2019.pdf	0,6948
10	PM Lingkungan Hidup No. 23 Tahun 2016.pdf	0,6887

Query 30 : Tata cara pembuangan limbah ke laut.

No	Dokumen	Entropy
1	PM Lingkungan Hidup No. 12 Tahun 2006.pdf	0,5680

2	PM Lingkungan Hidup No. 12 Tahun 2018.pdf	0,5632
3	PM Lingkungan Hidup No. 102 Tahun 2018.pdf	0,4879
4	KM Lingkungan Hidup No. 259 Tahun 2018.pdf	0,4732
5	PM Lingkungan Hidup No. 20 Tahun 2020.pdf	0,4512
6	PP No. 19 Tahun 1999.pdf	0,4480
7	PM Lingkungan Hidup No. 54 Tahun 2017.pdf	0,4000
8	UU No. 18 Tahun 2008.pdf	0,3508
9	PM Lingkungan Hidup No. 5 Tahun 2010.pdf	0,3483
10	PM Lingkungan Hidup No. 6 Tahun 2010.pdf	0,3404