

Terbit online pada laman web jurnal: <http://jurnal.iaii.or.id>

JURNAL RESTI

(Rekayasa Sistem dan Teknologi Informasi)

Vol. 5 No. 1 (2021) 63 - 68

ISSN Media Elektronik: 2580-0760

Temu Kembali Informasi pada Soal Ujian dengan Rencana Pembelajaran Menggunakan *Vector Space Model*

Amalia Beladinna Arifa¹, Gita Fadila Fitriana², Ananda Rifkiy Hasan³

¹Prodi S1 Teknik Informatika, Fakultas Informatika, Institut Teknologi Telkom Purwokerto

^{2,3}Prodi S1 Rekayasa Perangkat Lunak, Fakultas Informatika, Institut Teknologi Telkom Purwokerto

¹amalia@ittelkom-pwt.ac.id, ²gita@ittelkom-pwt.ac.id, ³18104004@ittelkom-pwt.ac.id

Abstract

One way to find out the quality of exam questions is by looking at the rules for writing exam questions made based on the subject or discussion contained in the learning plan document. Therefore, the exam questions that are arranged must be adjusted to the main material in each subject learning achievement. This study discusses the implementation of the concept in information retrieval systems using the Vector Space Model method. The Vector Space Model method has an advantage in query matching because it is able to match only part of the query with existing documents. In addition, the Vector Space Model method is also easy to adapt by adjusting parameters, including weighting parameters. The weighting calculation for each term that appears in the document uses TF-IDF. The purpose of this study is to design an information retrieval system to find the suitability of the exam question query with the subject contained in the learning plan document. The suitability is sorted based on the similarity value of the calculation results, from the largest value to the smallest value in the form of a percentage.

Keywords: information retrieval, search engine, exam questions, learning plan document, vector space model.

Abstrak

Salah satu cara untuk mengetahui kualitas soal ujian yaitu dengan melihat kaidah penulisan soal ujian yang dibuat berdasarkan materi pokok atau bahasan yang terdapat di dokumen rencana pembelajaran. Soal ujian yang disusun harus disesuaikan dengan materi pokok pada setiap capaian pembelajaran mata kuliah. Penelitian ini membahas mengenai penerapan konsep dalam sistem temu kembali informasi menggunakan pendekatan metode *Vector Space Model* (VSM). Metode *Vector Space Model* memiliki kelebihan dalam pencocokan *query* karena mampu mencocokkan hanya sebagian *query* saja dengan dokumen yang ada. Selain itu metode *Vector Space Model* juga mudah beradaptasi melalui penyesuaian parameter, termasuk parameter pembobotannya. Perhitungan pembobotan tiap kata (*term*) yang muncul pada dokumen menggunakan TF-IDF. Implementasi sistem temu kembali informasi (*retrieval of information*) sebagai media pencari kecocokan dan kesesuaian antara *query* soal ujian dengan materi pokok dokumen Rencana Pembelajaran (RPS), hal tersebut merupakan tujuan dari penelitian ini. Kesesuaian tersebut diurutkan berdasarkan nilai kemiripan hasil perhitungannya, dari nilai terbesar ke nilai terkecil dalam bentuk persentase.

Kata kunci: temu kembali informasi, Rencana Pembelajaran Semester, *vector space model*.

1. Pendahuluan

Pendidikan di setiap jenjang diselenggarakan berdasarkan kurikulum yang ditetapkan dalam standar isi pendidikan nasional. Peraturan kememenristekdikti No. 44 Tahun 2015 tentang Standar Nasional Pendidikan Tinggi, kurikulum dimaknai sebagai

“Seperangkat rencana dan pengaturan mengenai capaian pembelajaran lulusan, bahan kajian, proses dan penilaian yang digunakan sebagai pedoman penyelenggaraan program studi” [1]. Kurikulum pendidikan tinggi kemudian dikembangkan lagi menjadi Rencana Pembelajaran Semester (RPS). RPS

merupakan gambaran prosedur pembelajaran yang dirancang untuk memenuhi capaian pembelajaran lulusan yang dibebankan di suatu mata kuliah. Dosen diharuskan untuk menyusun dokumen RPS sebelum memulai proses pembelajaran untuk setiap mata kuliah yang diampunya. Isi dari dokumen tersebut setidaknya terdiri dari peta capaian pembelajaran mata kuliah, deskripsi mata kuliah, tujuan perkuliahan, capaian pembelajaran lulusan yang dibebankan di mata kuliah, capaian pembelajaran mata kuliah, materi pokok, strategi pembelajaran, evaluasi/penilaian pembelajaran, daftar rujukan, serta jadwal perkuliahan.

Proses pembelajaran terdapat evaluasi yang merupakan faktor penting. Dikarenakan posisi evaluasi setara dengan kesesuaian tujuan pembelajaran [2]. Salah satu bagian dari kegiatan evaluasi adalah pembuatan instrumen penilaian yang digunakan untuk mengukur kemampuan dan hasil belajar mahasiswa sesuai dengan capaian pembelajaran mata kuliah. Salah satu bentuk instrumen penilaian adalah soal ujian atau tes. Soal ujian sebagai instrumen penilaian adalah sekumpulan pertanyaan yang diberikan kepada siswa untuk mendapatkan jawaban dari siswa dalam bentuk lisan, tulisan atau perbuatan [3]. Soal ujian dalam bentuk tulisan sering digunakan untuk mengukur kemampuan mahasiswa dalam memahami materi yang telah diberikan oleh dosen, khususnya ketika di pertengahan pembelajaran (UTS) dan di akhir pembelajaran (UAS).

Dosen diharapkan memiliki kemampuan yang memadai dalam mengembangkan dan menyusun soal ujian untuk mengukur hasil belajar mahasiswa [4]. Kemampuan dosen dalam mengembangkan soal ujian dapat mempengaruhi kualitas belajar mahasiswa. Jika dosen tidak mampu dalam membuat soal ujian yang baik, maka mahasiswa mengalami kesulitan dalam pemahaman yang tidak tergalai sepenuhnya. Sehingga hal ini menghasilkan mutu lulusan menjadi rendah [5]. Salah satu cara untuk mengetahui kualitas soal ujian yang dibuat yaitu dengan melihat kaidah penulisan soal ujian dari segi materi pokok yang terdapat di dalam dokumen rencana pembelajaran. Maka dari itu, soal ujian yang disusun harus disesuaikan dengan materi pokok pada setiap capaian pembelajaran mata kuliah.

Solusi yang diberikan yaitu menerapkan konsep temu kembali informasi sebagai sistem/mesin pencarian kesesuaian antara soal ujian dengan dokumen rencana pembelajaran berdasarkan pada materi pokok. Sistem pencarian digunakan untuk menemukan informasi yang relevan dengan kebutuhan *user* secara otomatis [6]. Teknik yang dapat digunakan untuk menyelesaikan permasalahan tersebut adalah *information retrieval* dan *text mining*. *Text mining* adalah proses menggali informasi dimana *user* berinteraksi dengan beberapa dokumen menggunakan suatu *tools analysis*, yang merupakan komponen-komponen dalam penambahan data [7]. Salah satu penelitian tentang *text mining* digunakan untuk pencari dan pencocokan dokumen teks

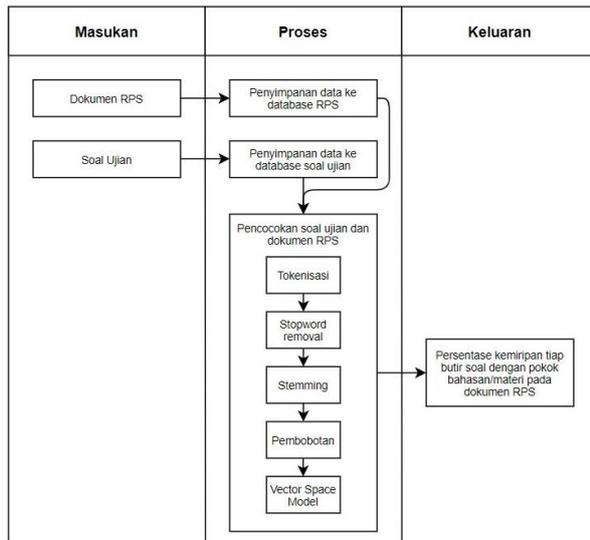
berbahasa Indonesia menggunakan *Vector Space Retrieval Model* [8]. Hasil dari penelitian tersebut menunjukkan bahwa sistem temu kembali informasi menggunakan metode *Vector Space Retrieval Model* dapat memberikan solusi pada mesin pencarian berupa informasi kecocokan teks dalam *database* menggunakan kata kunci tertentu. Di bidang pendidikan, metode *Vector Space Model* juga diimplementasikan pada aplikasi penentuan dosen penguji skripsi [9]. Hasil penelitian diperoleh akurasi sebesar 93,22%, dimana sistem dapat memberikan rekomendasi sebanyak tiga dosen penguji skripsi dengan syarat minimal 1 satu orang dosen memiliki kompetensi yang sama dengan topik judul skripsi.

Metode *Vector Space Model*, sebuah dokumen dipandang sebagai sebuah vektor yang memiliki jarak (*magnitude*) dan arah (*direction*) [10]. Metode *Vector Space Model* dapat dihasilkan tingkat kemiripan antar kata, frase dan dokumen dengan cara pembobotan kata (*term*). Metode *Vector Space Model* memiliki kelebihan dalam pencocokan *query* karena mampu mencocokkan hanya sebagian *query* dengan dokumen yang ada. Selain itu metode *Vector Space Model* juga mudah beradaptasi melalui penyesuaian parameter, termasuk parameter pembobotannya.

Tujuan dari penelitian ini yaitu implementasi sistem temu kembali informasi untuk mencari kesesuaian dan kecocokan antara soal dengan dokumen Rencana Pembelajaran Semester (RPS). Kesesuaian tersebut dihasilkan dari perhitungan menggunakan metode *Vector Space Model*, dimana beberapa penelitian sebelumnya telah berhasil menerapkan metode tersebut dalam pencarian kecocokan antar dokumen [11], [12], [13], [14].

2. Metode Penelitian

Proses pencarian kemiripan dalam sistem temu kembali informasi ini diterapkan menggunakan sistem penyimpanan dokumen Rencana Pembelajaran Semester (RPS) yang berformat (.xlsx). Proses pada *user* memasukkan *query* butir soal ujian secara manual ke dalam sistem. Penelitian ini bertujuan untuk mengetahui kemiripan tertinggi setiap butir soal ujian dengan materi pokok yang ada pada dokumen RPS. Dokumen akan dilakukan *indexing* yang berdasarkan pada semua isi dokumen. Selanjutnya, akan dilakukan *text-processing* yang meliputi *tokenizing*, *stopword removal*, dan *stemming*. Proses pembobotan pada metode ini menggunakan TF-IDF dan hasil yang diperoleh digunakan untuk perhitungan menggunakan *Vector Space Model* (VSM). Hasil akhir yang didapatkan adalah nilai kemiripan antara materi pokok pada dokumen RPS yang telah diurutkan dari yang paling tinggi nilainya berdasarkan *query* butir soal ujian yang dimasukkan. Kerangka aplikasi sistem pada penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Kerangka Aplikasi Sistem

2.1. Pre-processing

Tujuan dari dilakukannya tahap *pre-processing* pada *text-mining* adalah ketika dokumen atau teks banyak mengandung unsur yang tidak baku, terdapat imbuhan, simbol, angka dan sebagainya. Setiap teks akan dipecah menjadi struktur bagian kecil, biasanya menjadi suatu kata, yang akan memiliki makna yang sempit. Tahapan yang terdapat pada *pre-processing* yaitu:

a) Tokenizing

Tokenizing merupakan proses pemecahan sebuah dokumen menjadi unsur bagian terkecil yang disebut dengan token. Pemecahan ini dapat berupa pemecahan kalimat menjadi kata, simbol, frase dan entitas penting lainnya. Proses ini pada waktu yang bersamaan dilakukan juga penghapusan karakter tertentu, misalkan seperti simbol maupun tanda baca dan mengubah semua huruf menjadi huruf kecil (*lowercase*).

b) Stopword Removal

Kata-kata yang tidak memiliki arti dilakukan proses penghapusan kata untuk dijadikan sebagai kata kunci dalam pencarian dokumen. Kata-kata (*stopword*) yang diabaikan untuk kemudian dihapuskan, disimpan ke dalam *stoplists*. Tujuannya adalah untuk mengurangi jumlah kata pada *dataset* yang akan berpengaruh pada kecepatan dan performa kinerja sistem.

c) Stemming

Stemming merupakan proses pengubahan sebuah kata menjadi kata dasarnya dengan cara menghilangkan semua imbuhan kata. Imbuhan kata tersebut meliputi awalan (*prefixes*), sisipan (*infixes*), akhiran (*suffixes*), serta awalan dan akhiran (*confixes*) pada kata turunannya. *Library* yang digunakan untuk melakukan proses *stemming* pada penelitian ini adalah Sastrawi 1.0.1.

2.2. Perhitungan Menggunakan *Vector Space Model*

Setelah data melalui tahap *pre-processing*, akan dilakukan proses perhitungan kemiripan antar *query* dan dokumen menggunakan metode *Vector Space Model*. Proses menghitung kemiripan diperlukan suatu pembobotan terlebih dahulu. *Term Frequency – Inverse Document Frequency* (TF-IDF) merupakan faktor untuk menghitung bobot pada pengambilan suatu informasi. Data yang telah melalui tahap *pre-processing* memiliki bentuk numerik. TF-IDF digunakan untuk mengubah data yang berupa kata-kata menjadi numerik. Nilai TF diperoleh dari banyaknya kemunculan kata (*term*) pada sebuah dokumen. Nilai TF akan besar jika kata tersebut sering muncul dan akan kecil jika kata tersebut jarang muncul. Nilai IDF diperoleh dari banyaknya kata (*term*) yang dicari pada koleksi dokumen yang ada pada *database*.

Penelitian ini, nilai TF_i dihitung menggunakan nilai frekuensi kemunculan setiap kata. Nilai IDF_i dihitung menggunakan persamaan:

$$IDF_i = \log_{10} \left(\frac{N}{1+DF_i} \right) \quad (1)$$

dengan N adalah jumlah dokumen dan DF_i adalah jumlah dokumen dimana kata (*term*) muncul di dalamnya.

Perhitungan TF-IDF menggunakan persamaan:

$$W_{ij} = TF_{ij} \times \log_{10} \left(\frac{N}{1+DF_i} \right) \quad (2)$$

dengan W_{ij} adalah bobot dokumen, TF_{ij} adalah frekuensi kemunculan kata (T_i) pada dokumen (D_j), N adalah jumlah dokumen dan DF_i adalah jumlah dokumen dimana kata (T_i) muncul di dalamnya.

Hasil dari persamaan (1) dan (2) akan dimasukkan ke dalam suatu matriks dengan cara mengkalikan nilai TF_{ij} dan nilai IDF_i . Hasil perkalian tersebut kemudian akan dimasukkan ke dalam suatu dokumen baru. Dokumen ini akan diberikan *index* sesuai dengan kandidat kata kunci. Setiap kandidat kata kunci akan memiliki bobot, dimana bobot ini berasal dari perkalian antara vektor kata pada *query* (W_{iq}) dengan vektor dokumen (W_{ij}).

Perhitungan jarak pada *query* menggunakan persamaan:

$$|q| = \sqrt{\sum_{i=1}^t (W_{iq})^2} \quad (3)$$

dengan $|q|$ adalah jarak *query* dan W_{iq} adalah bobot *query* yang dihitung dari hasil pengindeksan dan pembobotan kata (*term*) dari *query*.

Sedangkan perhitungan jarak pada dokumen menggunakan persamaan:

$$|d_j| = \sqrt{\sum_{i=1}^t (W_{ij})^2} \quad (4)$$

dengan $|d_j|$ adalah jarak dokumen dan W_{ij} adalah bobot dokumen yang dihitung dari hasil pengindeksan dan pembobotan kata (*term*) dari kandidat kata kunci.

Perhitungan kemiripan antara *query* dan dokumen dilakukan dengan perkalian antara vektor *query* dan vektor kandidat kata kunci pada dokumen kemudian dibagi dengan perkalian antara nilai absolut jarak *query* dan jarak dokumen, menggunakan persamaan:

$$Sim(q, d_j) = \frac{q \times d_j}{|q| \times |d_j|} = \frac{\sum_{i=1}^t W_{iq} \times W_{ij}}{\sqrt{\sum_{i=1}^t (W_{iq})^2 \times \sum_{i=1}^t (W_{ij})^2}} \quad (5)$$

dengan $|q|$ adalah jarak *query*, $|d_j|$ adalah jarak dokumen, W_{iq} adalah bobot *query* yang dihitung dari hasil pengindeksan dan pembobotan kata (*term*) dari *query*, dan W_{ij} adalah bobot dokumen yang dihitung dari hasil pengindeksan dan pembobotan kata (*term*) dari kandidat kata kunci.

Nilai kemiripan yang paling tinggi antara *query* butir soal ujian dengan materi pokok pada dokumen RPS akan diberikan oleh sistem dalam bentuk persentase.

3. Hasil dan Pembahasan

3.1. Pre-processing

Sebelum melakukan pencarian kemiripan antara *query* soal ujian dengan dokumen RPS, terlebih dahulu dilakukan proses pembersihan pada data teks yang akan digunakan. Terdapat sebuah *query* yang diuji yaitu satu butir soal UTS mata kuliah Kalkulus yang berbunyi: “Hitunglah integral trigonometri dari fungsi berikut! $\int \sin^7 x dx$ ”. *Query* tersebut kemudian dicari kemiripannya dengan data uji. Penelitian ini menggunakan data uji sebanyak 2 (dua) data materi pokok yang terdapat pada dokumen RPS mata kuliah Kalkulus, yang dapat dilihat pada Tabel 1.

Tabel 1. Data Uji Penelitian

Dokumen RPS Mata Kuliah Kalkulus	
Materi Pokok	Keterangan
1	1. kontrak kuliah
	2. pengintegralan dengan substitusi
	3. integral trigonometri
	4. substitusi yang merasionalkan
	5. pengintegralan parsial
	6. pengintegralan fungsi rasional
2	1. integral tak wajar
	2. bentuk tak tentu jenis 0/0
	3. bentuk tak tentu yang lain
	4. integral tak wajar dengan batas pengintegralan tak hingga
	5. integral tak wajar dengan fungsi integran tak hingga di satu titik tertentu

Setiap data uji terlebih dahulu akan melalui tahapan *pre-processing* yang terdiri dari *tokenizing*, *stopword removal*, dan *stemming* untuk selanjutnya dilakukan proses pembobotan menggunakan TF-IDF serta perhitungan kemiripan menggunakan VSM.

a) Tokenizing

Hasil pada tahap ini adalah kalimat-kalimat pada materi pokok di dokumen RPS yang sudah terpecah menjadi kata-kata, seperti terlihat pada Tabel 2.

Tabel 2. Hasil *Tokenizing*

Dokumen RPS Mata Kuliah Kalkulus			
Materi Pokok (D _i)	Keterangan (<i>Term</i>)		
1	kontrak kuliah	integral trigonometri	pengintegralan parsial
	pengintegralan dengan substitusi	substitusi yang merasionalkan	pengintegralan fungsi rasional
	integral tak wajar bentuk tak tentu jenis 0/0	yang lain integral tak wajar dengan batas pengintegralan di	tak wajar dengan fungsi integran tak hingga
	bentuk tak tentu	integral	titik tertentu

b) Stopword Removal

Pada tahap ini yaitu menghapuskan kata-kata yang tidak memiliki arti atau tidak relevan, seperti terlihat pada Tabel 3.

Tabel 3. Hasil *Stopword Removal*

Dokumen RPS Mata Kuliah Kalkulus			
Materi Pokok (D _i)	Keterangan (<i>Term</i>)		
1	kontrak kuliah	trigonometri substitusi	pengintegralan fungsi rasional
	pengintegralan substitusi integral	merasionalkan pengintegralan parsial	
	integral tak wajar bentuk tak tentu jenis 0/0	lain integral tak wajar dengan batas pengintegralan di	fungsi integran tak hingga satu titik tertentu
	bentuk tak tentu	integral	wajar

c) Stemming

Hasil pada tahap ini adalah kata-kata yang sudah tidak memiliki imbuhan awal atau akhir serta menjadi kata-kata dasar, seperti ditunjukkan pada Tabel 4.

Tabel 4 terlihat bahwa terdapat 8 (delapan) *term* yang dihasilkan oleh proses *pre-processing* dari masing-masing materi pokok pada sebuah dokumen RPS. Delapan *term* yang terbentuk kemudian dilakukan proses pembobotan *term* menggunakan metode TF-IDF.

Tabel 4. Hasil *Stemming*

Dokumen RPS Mata Kuliah Kalkulus		
Materi Pokok (D _i)	Keterangan (<i>Term</i>)	
1	kontrak kuliah	trigonometri rasional parsial
	integral substitusi	fungsi
2	integral wajar	batas fungsi
	bentuk jenis	integran titik

3.2. Pembobotan Menggunakan TF-IDF

Penelitian ini, dilakukan pencarian kemiripan 1 (satu) butir soal UTS mata kuliah Kalkulus yang berbunyi: “Hitunglah integral trigonometri dari fungsi berikut! $\int \sin^7 x dx$ ” dengan data 2 (dua) materi pokok pada sebuah dokumen RPS. Hasil dari penggunaan persamaan (1) terlihat pada Tabel 5.

Tabel 5. Hasil Perhitungan TF-IDF

Materi Pokok (D _i)	<i>Term</i>	TF _i		DF	IDF _i
		RPS	Soal		
1	kontrak	1	0	1	0,4771
	kuliah	1	0	1	0,4771
	integral	4	1	2	0,1761
	substitusi	2	0	1	0,4771
	trigonometri	1	1	2	0,1761
	rasional	2	0	1	0,4771
	parsial	1	0	1	0,4771
	fungsi	1	1	2	0,1761
	hitung	0	1	1	0,4771
	sin	0	1	1	0,4771
2	dx	0	1	1	0,4771
	integral	4	1	2	0,1761
	wajar	3	0	1	0,4771
	bentuk	2	0	1	0,4771
	jenis	1	0	1	0,4771
	batas	1	0	1	0,4771
	fungsi	1	1	2	0,1761
	integran	1	0	1	0,4771
	titik	1	0	1	0,4771
	hitung	0	1	1	0,4771
trigonometri	sin	0	1	1	0,4771
	dx	0	1	1	0,4771

Tabel 5 terlihat bahwa kolom *Term* merupakan indeks kata kunci yang terpilih pada setiap dokumen serta kolom TF_i merupakan jumlah *term* yang muncul pada setiap *query* dan dokumen. Kolom DF pada *term* “kontrak” diperoleh dari frekuensi dokumen RPS yang mengandung *term* “kontrak” yaitu sebanyak 1 (satu) dokumen. Kolom IDF_i merupakan hasil dari perhitungan menggunakan persamaan (1).

Nilai TF dan IDF didapatkan kemudian dilanjutkan dengan perhitungan bobot untuk setiap *term*. Kolom W_{ij} yang terlihat pada Tabel 6 merupakan nilai bobot dari setiap *term*, yang dapat dipandang sebagai matriks.

Tabel 6 terlihat bahwa setiap *term* memiliki nilai bobot yang diperoleh dari jumlah *term* pada setiap dokumen dikalikan dengan nilai IDF. *Term* “kontrak” pada

dokumen RPS memiliki nilai bobot sebesar 0,143627809 yang diperoleh dari perkalian antara nilai TF dengan IDF-nya menggunakan persamaan (2).

Tabel 6. Hasil Perhitungan Bobot

Materi Pokok (D _i)	<i>Term</i>	W _{ij}	
		RPS	Soal
1	kontrak	0,143627809	0
	kuliah	0,143627809	0
	integral	0,123082508	0,053008751
	substitusi	0,227644692	0
	trigonometri	0,053008751	0,053008751
	rasional	0,227644692	0
	parsial	0,143627809	0
	fungsi	0,053008751	0,053008751
	hitung	0	0,143627809
	sin	0	0,143627809
2	dx	0	0,143627809
	integral	0,123082508	0,053008751
	wajar	0,287255618	0
	bentuk	0,227644692	0
	jenis	0,143627809	0
	batas	0,143627809	0
	fungsi	0,053008751	0,053008751
	integran	0,143627809	0
	titik	0,143627809	0
	hitung	0	0,143627809
trigonometri	sin	0	0,143627809
	dx	0	0,143627809

Nilai bobot dari masing-masing *term*, dilanjutkan dengan perhitungan menggunakan *Vector Space Model*. Nilai bobot tersebut digunakan sebagai titik koordinat pada ruang vektor yang merepresentasikan setiap *query* dan dokumen sebagai sebuah vektor. Hasil perhitungan jarak antar *query* dan dokumen dapat dilihat pada Tabel 7.

Tabel 7. Hasil Perhitungan Jarak

Materi Pokok (D _i)	<i>Term</i>	Nilai Jarak	
		RPS	Soal
1	kontrak	0,020628948	0
	kuliah	0,020628948	0
	integral	0,015149304	0,002809928
	substitusi	0,051822106	0
	trigonometri	0,002809928	0,002809928
	rasional	0,051822106	0
	parsial	0,020628948	0
	fungsi	0,002809928	0,002809928
	hitung	0	0,020628948
	sin	0	0,020628948
2	dx	0	0,020628948
	Total	0,431625084	0,265172822
	integral	0,123082508	0,053008751
	wajar	0,287255618	0
	bentuk	0,227644692	0
	jenis	0,143627809	0
	batas	0,143627809	0
	fungsi	0,053008751	0,053008751
	integran	0,143627809	0
	titik	0,143627809	0
hitung	trigonometri	0	0,143627809
	sin	0	0,143627809
	dx	0	0,143627809
	Total	0,484574987	0,296876482

Tabel 7 terlihat bahwa masing-masing *query* dan dokumen RPS memiliki nilai jarak “Total”. Nilai tersebut didapatkan dari perhitungan menggunakan persamaan (3) untuk *query* butir soal ujian dan persamaan (4) untuk materi pokok pada dokumen RPS, dengan rincian sebagai berikut:

$$\text{Materi pokok 1: } |q| = 0,265172822$$

$$|D_1| = 0,431625084$$

$$\text{Materi pokok 2: } |q| = 0,296876482$$

$$|D_2| = 0,484574987$$

Masing-masing didapatkan nilai jarak pada setiap *query* dan dokumen serta nilai jarak absolutnya, selanjutnya akan dicari nilai kemiripan *query* butir soal ujian dengan kata kunci pada materi pokok dokumen RPS menggunakan persamaan (5). Proses selanjutnya, perangkian sesuai dengan nilai yang terbesar ke nilai yang terkecil dalam bentuk persentase.

Tabel 8. Hasil Perhitungan Kemiripan

No. Urut	Materi Pokok	Persentase
1	D ₁	10,6%
2	D ₂	6,5%

Tabel 8 menunjukkan hasil perhitungan kemiripan antara *query* soal ujian yang berbunyi “Hitunglah integral trigonometri dari fungsi berikut! $\int \sin^7 x dx$ ” dengan 2 (dua) materi pokok pada dokumen RPS yang diuji. Hasil tingkat kemiripan antara materi pokok pertama (D₁) dalam dokumen RPS dengan *query* butir soal ujian yang diuji yaitu sebesar 10,6%. Materi pokok pertama (D₁) memiliki lebih banyak kata (*term*) yang relevan atau mirip dengan kata kunci yang terdapat pada *query* soal ujian, jika dibandingkan dengan materi pokok kedua (D₂).

4. Kesimpulan

Hasil penelitian menunjukkan bahwa metode *Vector Space Model* dapat diimplementasikan sebagai sistem pencarian dalam menemukan kesesuaian antara soal ujian dengan dokumen rencana pembelajaran. Pencarian kesesuaian tersebut diurutkan berdasarkan nilai kemiripan hasil perhitungannya, dari nilai terbesar ke nilai terkecil dalam bentuk persentase. Nilai yang dihasilkan oleh sistem memiliki arti bahwa *query* butir soal ujian tersebut telah memiliki kesesuaian dengan suatu materi pokok pada dokumen rencana pembelajaran.

Ucapan Terima Kasih

Tim penulis mengucapkan terima kasih kepada Direktorat Riset dan Pengabdian Masyarakat, Direktorat Jenderal Penguatan Riset dan Pengembangan, Kementerian Riset, Teknologi dan Pendidikan Tinggi (KEMENRISTEKDIKTI) yang telah mendanai penelitian ini melalui skema program Penelitian Dasar Pemula tahun 2020.

Daftar Rujukan

- [1] Sitepu, B. P. and Lestari, I., 2018. Pelaksanaan Rencana Pembelajaran Semester dalam Proses Pembelajaran di Perguruan Tinggi. *PERSPEKTIF Ilmu Pendidikan*, 32 (1), pp.43-51.
- [2] Sundari, D. and Pangaribuan, T. R., 2017. Analisis Soal Semester Ganjil Mata Pelajaran Bahasa Indonesia Kelas VII SMP Negeri 2 Percut Sei Tuan Tahun Pembelajaran 2016/2017. *Asas: Jurnal Sastra*, 6 (2), pp.1-10.
- [3] Sudjana, N., 2005. *Penilaian Hasil Proses Belajar Mengajar*. Bandung: PT. Remaja Rosdakarya.
- [4] Munadi, S., 2011. Analisis Validasi Kualitas Soal Tes Hasil Belajar pada Pelaksanaan Program Pembelajaran. *Jurnal Cakrawala Pendidikan*, 1 (1), pp.145-159.
- [5] Taib, E. N., 2014. Analisis Kualitas Aspek Materi Butir Soal Buatan Dosen. *Jurnal Biotik*, 2 (2), pp.116-121.
- [6] Sugara, B., Dody and Donny, 2019. Sistem Temu Kembali Informasi pada Gejala Autisme dengan Metode Vector Space Model. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 3 (2), pp.257-264.
- [7] Wisnu, D. and Hetami, A., 2015. Perancangan Information Retrieval (IR) untuk Pencarian Ide Pokok Teks Artikel Berbahasa Inggris dengan Pembobotan Vector Space Model. *Jurnal Ilmiah Teknologi dan Informasi*, 9 (9).
- [8] Karyono, G. and Utomo, F. S., 2012. Temu Balik Informasi pada Dokumen Teks Berbahasa Indonesia dengan Metode *Vector Space Model*. In: Seminar Nasional Teknologi Informasi & Komunikasi Terapan 2012 (SEMANTIK 2012), Semarang.
- [9] Siregar, R. R. A., Sinaga, F. A., and Arianto, R., 2017. Aplikasi Penentuan Dosen Penguji Skripsi Menggunakan Metode TF-IDF dan Vector Space Model. *Computatio: Journal of Computer Science and Information Systems*, 1 (2), pp.171-186.
- [10] Fauzi, A. and Ginabila, 2018. Information Retrieval System pada File Pencarian Dokumen Tesis Berbasis Text Menggunakan Metode Vector Space Model. *Jurnal PILAR Nusa Mandiri*, 14 (2), pp.41-46.
- [11] Lemman, D. and Andesa, K., 2017. Implementasi *Vector Space Model* Untuk Meningkatkan Kualitas Pada Sistem Pencarian Buku Perpustakaan. In: Seminar Nasional Informatika 2015 (SNIf), 1 (1), pp.8-15.
- [12] Gawande, P. S. and Suryawanshi, A., 2015. Improving Web Page Classification by Vector Space Model. *International Journal of Innovative Research in Computer and Communication Engineering*, 3 (2), pp.1252-1257.
- [13] Suprianto, Sunardi, and Fadlil, A., 2019. Aplikasi Sistem Temu Kembali Angket Mahasiswa Menggunakan Metode Generalized Vector Space Model. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, 6 (1), pp.33-40.
- [14] Jovita, Linda, Hartawan A. and Suhartono, D., 2015. Using Vector Space Model in Question Answering System. *Procedia Computer Science*, 59 (2015), pp.305-311.