

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB II

LANDASAN TEORI

2.1 Perpustakaan

Perpustakaan adalah sebuah organisasi, berupa lembaga atau unit kerja yang bertugas menghimpun koleksi pustaka dan menyediakan bagi penggunaannya untuk dimanfaatkan. Walaupun ada beberapa jenis perpustakaan seperti perpustakaan umum, perpustakaan khusus, perpustakaan sekolah, perpustakaan perguruan tinggi, dan perpustakaan nasional secara umum tujuannya adalah sama. Namun setiap jenis perpustakaan mempunyai tujuan tertentu. Oleh sebab itu organisasi untuk tiap-tiap jenis perpustakaan bisa berbeda, karena, secara khusus tujuannya tiap jenis perpustakaan tidak sama. (Widiasa , Ketut ;, 2007)

2.1.1 Tujuan perpustakaan

Perpustakaan secara umum memiliki tujuan untuk melakukan layanan informasi literer kepada masyarakat. Tujuan khusus dibedakan oleh jenis perpustakaan karena dari jenis perpustakaan memiliki perbedaan layanan bagi penggunaannya (user). Tujuan tersebut dijabarkan sebagai berikut . (Widiasa , Ketut ;, 2007):

a. Perpustakaan Nasional RI

Merupakan Perpustakaan Nasioal yang berkedudukan di Ibu Kota Negara Indonesia yang mempunyai jangkauan dan ruang lingkup secara Nasional dan merupakan salah satu Lembaga Pemerintah Non Departemen (LPND) yang bertanggung jawab kepada Presiden.

b. Badan Perpustakaan Daerah

Badan perpustakaan daerah atau lembaga lain yang sejenis adalah yang berkedudukan di tiap provinsi di Indonesia yang mengelola perpustakaan.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

c. Perpustakaan Umum

Perpustakaan umum diibaratkan sebagai Universitas Rakyat atau Universitas Masyarakat, maksudnya adalah bahwa perpustakaan umum merupakan lembaga pendidikan bagi masyarakat umum.

d. Perpustakaan Perguruan Tinggi

Perpustakaan yang berada di Perguruan Tinggi, baik berbentuk Universitas, Akademi, Sekolah Tinggi, ataupun Institut. Keberadaan, tugas dan fungsi perpustakaan tersebut adalah dalam rangka melaksanakan Tri Dharma Perguruan Tinggi, meliputi pendidikan, penelitian atau riset dan pengabdian kepada masyarakat.

e. Perpustakaan Sekolah

Perpustakaan sekolah berada di sekolah, dikelola sekolah, dan berfungsi untuk sarana kegiatan belajar mengajar, penelitian sederhana, menyediakan bahan bacaan, dan tempat rekreasi.

f. Perpustakaan Khusus

Perpustakaan khusus berada pada lembaga-lembaga pemerintahan dan swasta. Perpustakaan tersebut diadakan sebagai sumber informasi dan ilmu pengetahuan yang berkaitan baik langsung maupun tidak langsung dengan instansi induknya.

g. Perpustakaan Lembaga Keagamaan

Merupakan perpustakaan yang dimiliki dan dikelola oleh lembaga-lembaga keagamaan, misalnya perpustakaan, masjid, gereja.

h. Perpustakaan Internasional

Perpustakaan Internasional Merupakan perpustakaan internasional yang memiliki koleksi yang menyangkut negara-negara anggota atau negara-negara yang berafiliasi kepada lembaga dunia tersebut. Perpustakaan ini dikelola dan diselenggarakan lembaga internasional.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

i. Perpustakaan Kantor Perwakilan Negara-negara Asing

Merupakan perpustakaan yang dimiliki dan diselenggarakan oleh lembaga atau kantor perwakilan Negara masing-masing. Contohnya perpustakaan lembaga kebudayaan amerika dan pusat kebudayaan jepang

j. Perpustakaan Digital

Perpustakaan digital bukan merupakan salah satu jenis perpustakaan yang berdiri sendiri, tetapi merupakan pengembangan dalam system pengelolaan dan layanan perpustakaan.

2.1.2 Tugas pokok perpustakaan

Karena tujuan perpustakaan memberikan layanan informasi kepada pengguna perpustakaan, maka tugas pokok pustakawan adalah sebagai berikut : (Widiassa , Ketut ;, 2007)

- 1) Menghimpun bahan pustaka yang meliputi buku dan nonbuku sebagai sumber informasi
- 2) Mengolah dan merawat pustaka.
- 3) Memberi layanan bahan pustaka kepada pengguna pustaka.

2.2 *Information Retrieval*

Information Retrieval (IR) adalah sebuah proses pencarian atau menemukan kembali informasi yang dibutuhkan dari sebuah database atau sistem penyimpanan dan penelusuran informasi yang ada. Didalam sistem temu kembali informasi mensyaratkan ada kebutuhan suatu informasi dari pengguna, ada pencarian dokumen atau data yang berisi informasi yang dikelompokan atau diorganisasikan dalam sebuah sistem yang memudahkan pencarian temu kembali informasi dan strategi penelusuran yang pasti sehingga data yang dibutuhkan dapat ditemukan kembali. (Putung. 2016)

Sistem temu kembali informasi terbagi tiga komponen utama yaitu masukan (*input*), memproses (*processor*), dan keluaran (*output*). Menurut *manning'et'al* (2009) *information retrieval* (IR) adalah menemukan kembali

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

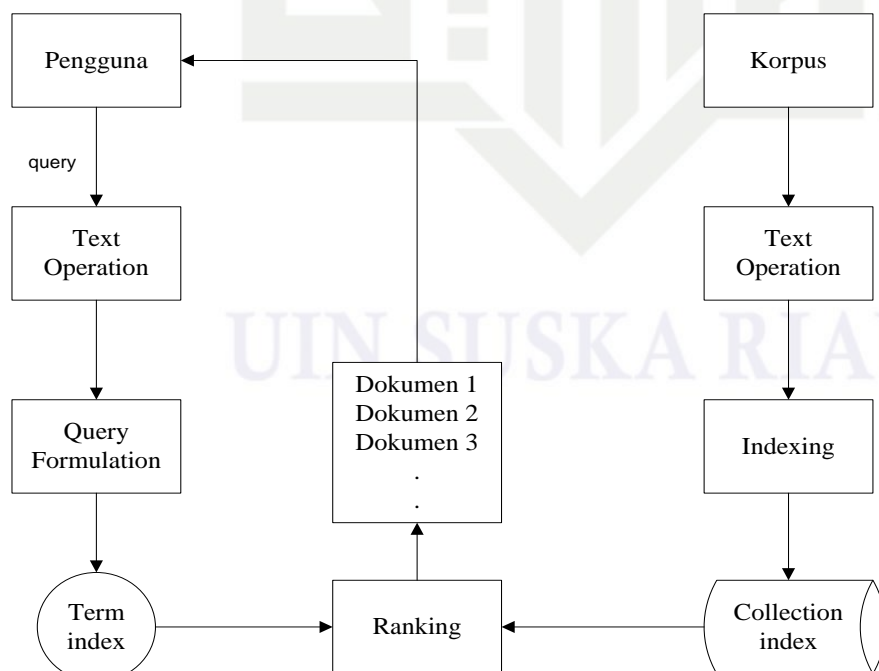
bahan (dokumen) yang bersifat tidak terstruktur (*teks*) yang memenuhi kebutuhan informasi dari dalam koleksi besar (database). (Putung. 2016)

Information Retrieval (IR) terbagi dari beberapa bagian yang dijabarkan sebagai berikut :

- a. *Text Operations* yaitu meliputi pemilihan kata-kata dalam query maupun dokumen (*term selection*) dalam proses transformasi dokumen atau *query* menjadi *term index*.
- b. *Query formulation* yaitu memberi bobot pada kata *query* yang terindeks
- c. Ranking yaitu mencari dokumen-dokumen yang relevan terhadap *query* dan mengurutkan dokumen tersebut sesuai dengan *query*.
- d. *Indexing* yaitu proses awal membangun basis data indeks dari koleksi dokumen terlebih dahulu sebelum pencarian dokumen dilakukan. (Putung. 2016)

2.3 Arsitektur *Information Retrieval*

Arsitektur sebuah *information retrieval* (IR) dapat kita pahami dengan gambar dibawah ini : (Putung. 2016)



Gambar 2.1. Arsitektur *Information Retrieval System* (IRS)



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.4 Text Mining

Text mining didefinisikan sebagai proses dimana pengguna berinteraksi dengan koleksi dokumen dari waktu ke waktu menggunakan seperangkat alat analisis. *Text mining* berusaha mengekstrak informasi yang berguna melalui identifikasi dan eksplorasi pola yang menarik dari sumber data yang ada. *Text mining* memiliki tujuan yaitu mengumpulkan informasi yang berguna dari data yang ada, dimana data tersebut berupa teks dalam bahasa baku. Proses *text mining* yang khas meliputi kategorisasi *teks*, *text clustering*, ekstraksi konsep/entitas, sentiment analysis, penyimpulan dokumen, dan pemodelan relasi entitas (pembelajaran hubungan antara entitas bernama)

Menurut penelitian terdahulu, mengatakan pada penelitian *text mining* diperlukan tahapan *text preprocessing* pada koleksi dokumen lalu menyimpan informasi tersebut kedalam struktur data. Pendekatan *text mining* didasarkan pada pemikiran yaitu dokumen teks dapat diwakilkan oleh satu set kata, dimana dokumen teks digambarkan berdasarkan pada set kata-kata yang terkandung didalamnya (Olivita, Dhita. 2016).

2.5 Tahap-tahapan Text Mining

Pada proses *presprosesing* dimana deskripsi ditangani untuk dapat siap diproses memasuki tahap *Text Mining* diantara proses itu diantar lain yaitu sebagai berikut (B, DwijaWisnu; Hetami, Anandini;. 2015)

a. Tokenizing

Sebuah proses yang dilakukan untuk memotong tiap kata kalimat didalam sebuah kalimat menggunakan spasi sebagai delimiter sehingga menjadikan lebih bermakna atau menjadikan kata-kata.

b. Stopword

Filtering atau *stopwords removal* sebuah proses manyaring kata yang didapat dari proses *tokenizing* yang dianggap tidak terlalu penting atau tidak bermakna dalam proses *text mining* yang disebut *stoplist*. *Stoplist* atau *stopwords* berisi kata-kata umum yang sering muncul dalam sebuah dokumen dalam jumlah yang banyak namun tidak memiliki kaitan atau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

makna dengan tema tertentu. Setiap kata yang diproses dari tokenizing akan dicocokkan kedalam database, jika kata cocok akan diproses kedalam stopward maka kata tersebut akan dihilangkan, sementara yang tidak cocok akan diproses ketahap selanjutnya

c. *Stemming*

Stemming adalah proses konversi term ke bentuk dasarnya. Hal ini dilakukan dengan menghilangkan akhiran atau awalan dari sebuah kata. algoritma yang dikhususkan untuk *stemming* bahasa Indonesia antara lain, Algoritma Porter, Nazief dan Adriani, Vega, CS, ECS .

Metode *stemming* memerlukan masukan berupa kata yang terdapat dalam suatu dokumen, dengan menghasilkan keluaran berupa root word.

d. *Analyzing*

Keterhubungan antar kata dalam dokumen akan ditentukan dengan menghitung frekuensi term pada dokumen, atau merupakan tahap menentukan seberapa jauh keterhubungan antara kata-kata dengan dokumen yang ada. Tahap ini lebih sering dikenal dengan tahap pembobotan.

2.6 Termfrequency-inverse document frequency (TF-IDF)

Metode *Termfrequency-inverse document frequency* (TF-IDF) ini merupakan untuk menghitung bobot setiap kata pada dokumen yang paling umum digunakan pada *Information Retrieval* (IR). Metode ini juga dikenal efisien, mudah dan memiliki hasil yang akurat. Metode ini akan menghitung nilai *term frequency* (TF) dan *inverse document frequency* (IDF) pada setiap token (kata) di setiap dokumen dalam korpus. Metode ini akan menghitung bobot setiap kata t di dokumen d dengan rumus : (Maarif, Abdul)

$$W_{dt} = tf_{dt} * IDF \dots\dots\dots(2.1)$$

Dimana :

- d : dokumen ke- d
- t : kata ke-t dari kata kunci

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- W : bobot dokumen ke-d terhadap kata ke-t
- Tf : banyaknya kata yang dicari pada sebuah document
- Idf : nilai idf didapatkan dari log2 (D/df)
- D : total dokumen
- df : banyak dokumen yang akan dicari

Dalam pencarian dibutuhkan metode pembobotan *term* agar pencarian lebih mudah difokuskan. Pada model ruang vektor, pembobotan terhadap *term* dilakukan dengan mengalikan bobot *tf* dan bobot *idf*, dikenal dengan pembobotan *tf-idf*.

- a. *Term frequency (tf)* adalah jumlah kemunculan sebuah *term* pada sebuah dokumen. Jika sebuah *term i* sering muncul pada sebuah dokumen, maka *query* yang mengandung *term* harus mendapatkan dokumen tersebut. Nilai sebuah *tf* dihitung berdasarkan kemunculan *term* dalam dokumen.
- b. *Inverse Document Frequency (idf)* adalah jumlah dokumen yang mengandung sebuah *term* yang dicari dari kumpulan dokumen yang ada.

$$Idf(i) = \log \left(\frac{N}{Df(i)} \right) \dots \dots \dots (2.2)$$

Dimana:

N= jumlah seluruh dokumen pada koleksi dokumen

df(i) = jumlah dokumen koleksi yang mengandung *term i*

- c. Bobot *Term* Pembobotan *tf x idf* untuk sebuah *term i* untuk dokumen *j* didapatkan dari hasil perkalian *tf* dan *idf*.

$$w_{i,j} = tf_{i,j} \times idf_i \dots \dots \dots (2.3)$$

Setelah bobot (W) masing-masing dokumen diketahui, maka akan dilakukan proses pengurutan dimana semakin besar nilai W, maka akan semakin besar tingkat similaritas dokumen tersebut terhadap kata kunci, demikian sebaliknya. (Maarif, Abdul)

Hak Cipta Dilindungi Undang-Undang

2.7 Cosine Similarity

Cosine similarity adalah sebuah proses untuk mengukur kesamaan antara dua buah vector dalam sebuah ruangan dimensi dimana mendapatkan nilai cosinus sudut dari perkalian dua buah vector yang dibandingkan karena cosinus dari 0^0 adalah 1 dan kurang dari 1 untuk nilai sudut yang lain, maka nilai similarity dari buah vektor dikatakan mirip ketika nilai dari cosine similarity adalah 1.

Cosine similarity digunakan untuk dalam ruang positif, dimana hasilnya dibatasi antara nilai 0 dan 1. kalau nilainya 0 maka dokumen tersebut akan teridentifikasi mirip jika hasil 1 maka dokumen itu tidak teridentifikasi mirip perhatikan bahwa batas ini berlaku untuk sejumlah dimensi, dan proses ini paling sering digunakan didalam ruang positif dimensi tinggi. Misalnya didalam *information retrieval*, masing-masing kata (*term*) diasumsikan sebagai dimensi yang berbeda dan dokumen ditandai dengan vektor dimana nilai masing-masing dimensi sesuai dengan berapa istilah muncul dalam dokumen (Ariantini, Dewa Ayu Rai, 2016). Berikut ada rumus *cosine similarity* (dermawan, surya; santoso, edy; muflikhah, laili, 2018):

$$CS(b_1, b_2) = \frac{\sum_{t=1}^n W_{t,b1} W_{t,b2}}{\sqrt{\sum_{t=1}^n W_{t,b1}^2} \cdot \sqrt{\sum_{t=1}^n W_{t,b2}^2}} \dots\dots\dots(2.4)$$

Perhitungan skalar didapatkan dari hasil perkalian antara bobot *query* dan perkalian bobot tiap dokumen dengan rumus $WD * W_{di}$ atau sama dengan potongan rumus *Cosine Similarity* yaitu $\sum_{t=1}^n W_{t,b1} W_{t,b2} \dots\dots\dots(2.5)$

Sedangkan perhitungan panjang vektor dengan potongan rumus *Cosine Similarity* yaitu : $\sqrt{\sum_{t=1}^n W_{t,b1}^2}$ dan $\sqrt{\sum_{t=1}^n W_{t,b2}^2} \dots\dots\dots(2.6)$

2.8 Clustering K-Means

Algoritma *K-Means* salah satu kemampuan yang dimiliki oleh data mining adalah kemampuannya untuk melakukan proses pengklusteran atau pengelompokan pada suatu data. Algoritma ini mampu meminimalkan jarak antara data ke clusteringnya. Algoritma k-means pada dasarnya melakukan 2 proses

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

diantaranya proses pendekteksian lokasi pusat clustering dan proses pencarian anggota dari tiap-tiap clustering. Proses clustering dimulai dengan mengidentifikasi data yang akan dikluster, X_{ij} ($i=1, \dots, n$; $j=1, \dots, m$) dengan n adalah jumlah data yang akan dikluster dan m adalah jumlah variabel. Pada awal iterasi, pusat setiap kluster ditetapkan secara bebas, C_{kj} ($k=1, \dots, k$; $j=1, \dots, m$) kemudian dihitung jarak antara setiap data dengan setiap pusat kluster. Untuk melakukan perhitungan jarak data K_{e-1} (x_i) pada pusat kluster ke- k (c_k), diberi nama (d_{ik}) dapat digunakan formula Euclidean.

Proses algoritma *Clustering K-Means* dalam penelitian ini digunakan untuk menghitung kemiripan antara dokumen yang dihasilkan oleh *Sphinx Search* sesuai dengan *Query* pengguna dan dalam kelas yang sama terhadap dokumen lainnya untuk membentuk daftar *Related Search* yaitu berupa daftar buku yang mungkin dapat dijadikan refrensi oleh pengguna. (Indriani, Aida ;,2013). Proses dasar Algoritma *K-Means* dapat dilihat dibawah ini (qolbi, ahmad asif;, 2016):

1. Menentukan jumlah *cluster*
2. Menentukan nilai *centroid*

Dalam menentukan nilai *centroid* untuk awal iterasi, nilai awal *centroid* dilakukan secara acak. Sedangkan jika menentukan nilai *centroid* yang merupakan tahap iterasi, maka digunakan rumus sebagai berikut :

$$\mu_k = \frac{1}{N} \sum_{q=0}^{N_k} x_q \dots \dots \dots (2.7)$$

dimana:

μ_k = titik central dari *cluster* ke- K

N_k = banyaknya data pada *cluster* ke- K

X_q = data ke- q pada *cluster* ke- K

3. Menghitung jarak antara titik *centroid* dengan titik tiap objek

Untuk menghitung jarak tersebut dapat menggunakan *Euclidean Distance*, yaitu:

$$D = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \dots \dots \dots (2.8)$$

Dimana:

Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

$D = euclidian\ distance$
 $I = \text{banyak nya objek}$
 (x,y) merupakan koordinat objek dan
 (s,t) merupakan koordinat *centroid*.

4. Pengelompokan objek, Untuk menentukan anggota *cluster* adalah dengan memperhitungkan jarak *minimum* objek. Nilai yang diperoleh dalam keanggotaan data pada *distancematriks* adalah 0 atau 1, dimana nilai 1 untuk data yang dialokasikan ke *cluster* dan nilai 0 untuk data yang dialokasikan
5. Kembali ke tahap 2, lakukan perulangan hingga nilai *centroid* yang dihasilkan tetap dan anggota *cluster* tidak berpindah ke *cluster* lain. (Suryanto, Hizkia Juan; lukito, Yuan;)

2.9 Recall dan Precision

Tujuan uji *Recall* dan *Precision* adalah untuk mengukur kinerja sistem temu kembali. Nilai *Precision* adalah proporsi dokumen yang terambil oleh sistem adalah relevan. Menurut Salton dalam (Amin, 2012) nilai *Precision* adalah proporsi dokumen relevan yang terambil oleh sistem

Rumusan *precision* :

$$precision = \frac{TP}{TP+FP} \dots\dots\dots(2.9)$$

Dimana :

TP = *True Positif*

FP = *False Positif*

Recall mengevaluasi kemampuan sistem temu kembali informasi untuk menemukan semua *item* yang relevan dari dalam koleksi dokumen dan didefinisikan sebagai persentase dokumen yang relevan terhadap *query* pengguna dan yang diterima.

$$recall = \frac{TP}{TP+FN} \dots\dots\dots(2.10)$$

Dimana :

TP = *True Positif*

FN = *False Negatif*

2.10 Penelitian Terkait

Penelitian terkait yang berkaitan dengan penelitian ini dapat dilihat pada tabel dibawah ini :

NO	PENULIS	TAHUN	JUDUL	KESIMPULAN
	Surya dermawan,edy Santoso	2018	Penentuan Pemenang Tender Menggunakan Kombinasi K- Nearest Neighbor dan Cosine Similarity (Studi Kasus PT. Unichem Candi Indonesia)	Refrensi yang diambil dalam penelitian ini adalah rumus consine similarty
	Arie Lumenta	2016	Penerapan Sistem temu kembali informasi pada kumpulan dokumen skripsi	Berdasarkan hasil dari penelitian yang dilakukan sistem temu kembali dapat melakukan implementasi pengindeksan otomatis dengan menggunakan metode TF-IDF banyak data 100 dokumen dengan waktu 328.85 detik.
	Dhita Olivita	2016	Perbandingan klasifikasi tugas akhir mahasiswa jurusan teknik informatika menggunakan metode Naie Bayes Classifier dan K-Nearest Neighbor	Dalam penelitian ini penulis mengambil penjelasan tentang text mining yaitu <i>text mining</i> yaitu memiliki tujuan untuk mengumpulkan informasi yang berguna dari data yang ada, dimana data tersebut berupa teks dalam baku,proses <i>text mining</i> yang

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Diindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

				<p>khas meliputi kategorisasi <i>teks</i>, <i>text clustering</i>, ekstraksi konsep/entitas, sentiment analysis, penyimpulan dokumen, dan pemodelan relasi entitas (pembelajaran hubungan antara entitas bernama)</p>
4	Ogie Nurdiana	2016	Perbandingan Metode Cosine Similarity dengan Metode Jaccard Similarity pada aplikasi Pencarian Terjemahan AL-Qura'an dalam bahasa Indonesia`	<p>Dengan menerapkan algoritma KNN sebagai klasifikasi dimana semakin banyak dokumen yang diuji maka semakin banyak pula sumber yang digunakan untuk dibandingkan karena KNN mempunyai sifat <i>Self-Learning</i> berarti mencari tetangga yang paling dekat dengan sets yang akan di klasifikasi.</p>
5	Anindya Kharina Wardhani	2016	Implementasi Algoritma K-Means untuk pengelompokan penyakit pasien puskesmas kajian pekalongan	<p>Berdasarkan penelitian ini bahwa dirumuskan jumlah Cluster sebanyak 2 buah sesuai dengan definisi nilai k dengan jumlah Cluster akut 376 item, cluster tidak akut ada 624 item</p>

Hak Cipta Diindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

				total jumlah data adalah 1000.
	Ahmad Asif Qolbi	2016	Penerapan metode clustering K-Means terhadap Dosen berdasarkan publikasi Jurnal Nasional dan Internasional	Refrensi jurnal yang diambil dalam penelitian ini adalah langkah-langkah metode K-means
	Hizkia juaan Suryanto, yuaan lukito	2016	Indoor Positioning System dengan Algoritma K-Means dan KNN	Refrensi jurnal yang diambil dalam penelitian ini adalah metode K-means.
8	Dwija Wisnu B,Anandini Hetami	2015	Perancangan information Retrieval (IR) untuk mencari Ide pokok Teks Artikel bahasa inggris dengan pembobotan Vektor Space Model.	Dalam penelitian ini penulis mengambil penjelasan dan pengertian tentang tahapan Text Mining
9	Abdul Azis Maarif	2015	Penerapan Algoritma TF-IDF untuk pencarian Karya Ilmiah	Penerapan algoritma TF-IDF untuk pencarian karya ilmiah dalam penelitian ini dapat menghasilkan karya ilmiah yang dapat dilihat atau diunduh dalam bentuk format PDF
10	Yoseph Samuel	2014	Implementasi Metode K-NN dengan Decision Rule untuk klasifikasi Suptopik Berita	Pada penelitian ini akan melakukan penggantian penggunaan majority menjadi decision rule dengan harapan agar pengguna algoritma

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

				K-NN dapat dimaksimalkan.
	Aida Indriani	2013	WAK-NN dan MST untuk information Retrieval system perpustakaan STMIK PPKIA Tarakanita Rahmawati Tarakan	Penelitian ini menerapkan WAK-NN sebagai klasifikasi data untuk sistem informasi perpustakaan. Proses klasifikasi terhadap 500 data membagi menjadi dua yaitu data latih dan data uji. Model klasifikasi dalam data uji diperoleh dari WAK-NN untuk data latih sebanyak 350 data. Dengan menggunakan model klasifikasi dari data latih, data uji sebanyak 150 data menghasilkan presentase akurasi kecocokan dihasilkan WAK-NN sebesar 81% waktu proses ±11 menit/ buku
	Fatkhul amin	2012	Dokumen abstraksi mahasissea fakultas ekonomi universitas indonesia dengan metode <i>vector space model</i>	Tokenisasi mampu dilewati dengan waktu komputasi 4 menit 41 detik. Menampilkan hasil pencarian komputasi rata-rata 1,5 detik
	Tri Susilowati	2012	Analisa Penerapan Customer Relationship Management (CRM) pada perpustakaan STMIK Pringsewu	Berdasarkan pembahasan dengan adanya konsep CRM penilaian terhadap mahasiswa pada

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

			dalam upaya meningkatkan kepuasan mahasiswa	perpustakaan STMIK Pringsewu yang merasa puas terhadap layanan (82,38%) dan yang tidak puas hanya (17,63%)
14	Neneng Komariah	2008	Aplikasi <i>Customer Relationship Management (CRM)</i> dalam layanan informasi diperpustakaan	Menjelaskan bagaimana menerapkan konsep (CRM) pada layanan informasi mulai dari pengadaan, pemrosesan, layanan informasi, dll. Dimana CRM merupakan strategi pemasaran yang fokus pada Stakeholder perusahaan yaitu pelanggan, dalam hal perpustakaan dimana yang menjadi pelanggan yaitu mereka yang sudah biasa menggunakan layanan perpustakaan
15	Ketut Widiasta	2007	Manajemen Perpustakaan Sekolah	Dalam penelitian ini menjelaskan tentang manajemen perpustakaan dan tugas pokok yang harus dilakukan perpustakaan