

PENGEMBANGAN MODEL SISTEM BERBASIS PENGETAHUAN PADA LEMBAGA KONSULTASI MANAJEMEN USAHA MIKRO KECIL DAN MENENGAH (UMKM)

Faiza Renaldi¹, Tacbir Hendro Pudjiantoro², Aceng Nursamsudin³

^{1,2,3}Program Studi Teknik Informatika Universitas Jenderal Achmad Yani, Cimahi

¹faiza@informatikaunjani.org, ²tacbir23501027@yahoo.cm, ³aceng.nursamsudin@gmail.com

Abstrak

Lembaga Konsultasi Manajemen (LKM) memiliki tujuan untuk membantu pelaku usaha dalam aktifitas bisnis mereka. Produk akhir yang diberikan oleh LKM adalah sebuah dokumen saran yang harus dilakukan oleh perusahaan dalam menghadapi permasalahan-permasalahan yang ada. Tentu saja, proses bantuan (asistensi) yang diberikan harus melalui sebuah tahap observasi awal dengan mengumpulkan data-data awal dan dibandingkan dengan isu-isu (permasalahan) apa saja yang terjadi di dalam perusahaan tersebut.

Di dalam penelitian ini akan dijabarkan pengembangan model Sistem Berbasis Pengetahuan yang menyimpan informasi berupa laporan-laporan dari hasil proses konsultasi antara LKM dengan perusahaan. Sistem ini memungkinkan pengguna (konsultan di dalam LKM) untuk mengakses kumpulan pengetahuan dan memilih pengetahuan mana yang relevan (sesuai dengan tipe industri, tipe masalah, dsb) untuk dapat digunakan pada konsultasi selanjutnya. Juga, terdapat fasilitas pembelajaran tak terselia untuk menghitung korelasi antar dokumen dengan referensi yang tersedia. Penelitian ini menggunakan studi kasus dan data laporan pada proyek Usaha Mikro Kecil dan Menengah (UMKM) yang dimiliki oleh LKM. Pengambilan studi kasus pada UMKM untuk Sistem Berbasis Pengetahuan ini memiliki tujuan terpenting untuk dapat menurunkan beban biaya observasi awal yang dimiliki oleh LKM sehingga dapat menurunkan total beban biaya yang akan dibebankan kepada UMKM dalam penggunaan jasa LKM.

Kata kunci : Sistem Berbasis Pengetahuan, Penggalan Teks, Pembelajaran Taktersedia, UMKM, Lembaga Konsultasi Manajemen.

Abstract

Management Consulting Institute (MCI) has the aim to assist SME (Small and Medium Enterprise) entrepreneurs in realizing their goals through their business activities. The final product provided by the MCI is suggestion of documents that must be done by the company in dealing with their problems. Of course, the processes are given through a stage of preliminary observations by collecting initial data and compared with the issues (problems) happened inside the company.

In this study, we will develop a Knowledge-Based Systems that stores information in the form of reports from the consultation process between the MCI and SMEs. The system allows users (consultants within MCIs) to access a collection of knowledge and pick the most (or a set of) relevant knowledge (according to industry type, the type of problem, etc.) to be used in subsequent consultations. Also, there is a non-supervised learning facility to calculate the correlation between the documents with references available.

This research uses case studies and report data on the projects Micro, Small and Medium Enterprises (SMEs) owned by MCI. Taking a case study on SMEs for Knowledge-Based Systems has the most important goal to reduce the burden of the cost during the preliminary observations which will lower the total cost will be charged to SMEs in the use of MCI services.

Keywords: Knowledge-Based Systems, Excavation Text, Not supervised Learning, MSME, Institute of Management Consultancy

1. Pendahuluan

Lembaga Konsultasi Manajemen (LKM) adalah lembaga yang memiliki tujuan untuk membantu pelaku usaha dalam aktivitas bisnis mereka. Produk akhir yang diberikan oleh LKM adalah sebuah dokumen saran yang harus dilakukan oleh perusahaan dalam menghadapi permasalahan-permasalahan yang ada.

Proses asistensi yang dilakukan oleh lembaga memberikan aset pengetahuan tersendiri pada individu konsultan yang melakukan proyek konsultasi. Namun sayangnya, aset pengetahuan hasil dari proses asistensi tersebut tersebar dan tersimpan pada masing-masing individu konsultan, sehingga aset pengetahuan tersebut tidak diketahui oleh lembaga secara pasti. Akibatnya, proses asistensi yang dilakukan menjadi kurang efektif, karena tahap observasi data dalam mencari data-data awal yang dibutuhkan untuk proses tahap observasi berikutnya akan dilakukan secara berulang. Pengulangan tersebut terjadi karena proses transfer pengetahuan antar konsultan sangat sedikit, terjadi hanya pada saat para konsultan tersebut baru bergabung dengan lembaga, dan jika terjadi diskusi antar sesama konsultan.

Kondisi pengulangan pada tahap observasi data sebenarnya masih dapat dihindari dengan digunakannya dokumen-dokumen hasil proyek konsultasi sebelumnya, yaitu dengan cara menggali informasi dan pengetahuan yang telah terdokumentasikan pada dokumen-dokumen tersebut. Namun sayangnya, akses yang dilakukan untuk mendapatkan informasi dan pengetahuan yang tersimpan pada dokumen-dokumen tersebut akan cukup sulit untuk dilakukan. Hal ini dikarenakan semakin bertambahnya jumlah dari dokumen-dokumen hasil proyek konsultasi, sehingga untuk menggali informasi dan pengetahuan yang tersimpan pada dokumen-dokumen tersebut pun akan memakan waktu yang cukup lama.

Permasalahan atas akses terhadap informasi dan pengetahuan yang tersimpan pada dokumen-dokumen hasil proyek konsultasi, dapat terbantu dengan membangun sebuah sistem (Knowledge Management System).

Pada implementasinya aset pengetahuan yang tersimpan pada isi dokumen-dokumen hasil proyek konsultasi harus dapat di ekstraksi ke dalam sistem, sehingga informasi dan pengetahuan tersebut dapat dengan mudah untuk diperoleh kembali oleh pengguna. Proses ekstraksi isi dokumen ke dalam sistem dapat dilakukan dengan berbagai cara, salah satunya dengan menggunakan metode/teknik pada text mining.

Text mining merupakan salah satu bidang khusus dari data mining. Text mining dapat digunakan untuk mendapatkan informasi yang berguna dari sekumpulan dokumen. Dalam memberikan solusi, text mining mengadopsi dan mengembangkan teknik dari bidang lain, seperti data mining, information retrieval, machine learning, dan natural language processing, dsb.

1.1 Rumusan Masalah

Berdasarkan pada latar belakang masalah yang telah diuraikan, dapat dirumuskan permasalahan yang akan dibahas dalam penelitian ini adalah bagaimana untuk mengekstraksi data pada dokumen-dokumen hasil proyek konsultasi antara lembaga dengan perusahaan dengan menggunakan algoritma text mining, sehingga informasi dan pengetahuan yang tersimpan pada dokumen-dokumen tersebut dapat diperoleh dengan cepat, tepat, dan mudah untuk digunakan kembali.

1.2 Batasan Masalah

Adapun yang menjadi pembatasan masalah pada penelitian tugas akhir ini adalah :

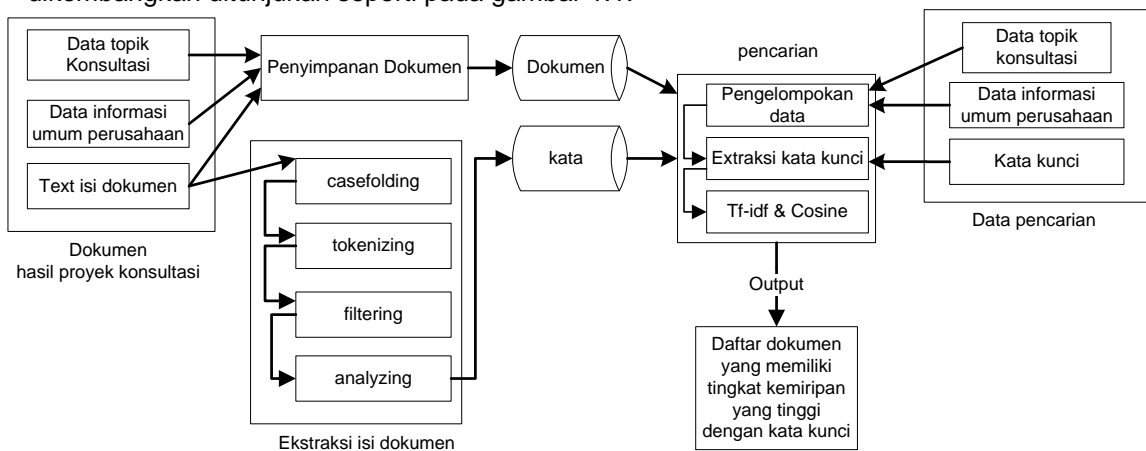
1. Sistem yang dibangun ditujukan untuk mengelola laporan hasil proyek konsultasi antara Lembaga Konsultasi Manajemen (LKM) dengan perusahaan.
2. Proses text mining yang dilakukan untuk melakukan ekstraksi data pada sistem hanya meliputi tahapan casefolding, tokenizing, filtering, dan analyzing.
3. Informasi dan pengetahuan awal untuk sampel data yang akan dikelola oleh sistem berasal dari informasi dan pengetahuan yang digunakan untuk menangani permasalahan pada Usaha Mikro Kecil dan Menengah (UMKM) yang ada pada internet, dan hanya mencakup UMKM yang berada pada provinsi Jawa Barat.

1.3 Tujuan dan Kegunaan Penelitian

Adapun yang menjadi tujuan dan kegunaan dari penelitian tugas akhir ini adalah untuk membangun sistem yang dapat digunakan untuk melakukan penyimpanan, pengelolaan, dan pengintegrasian informasi dan pengetahuan yang terdapat pada dokumen-dokumen hasil proyek konsultasi, dan mengetahui penerapan teknik / metode dari text mining yang dapat digunakan untuk menggali informasi dari sekumpulan dokumen, sehingga informasi yang tersimpan dapat digunakan kembali dengan cepat dan mudah, dan untuk meminimalisir hilangnya aset pengetahuan dari organisasi.

2. Metode Penelitian

Sistem yang akan dikembangkan didasarkan pada kebutuhan lembaga untuk mengakses informasi dan pengetahuan dari laporan-laporan hasil proyek konsultasi sebelumnya, sehingga sistem baru yang akan dikembangkan, dapat memungkinkan pengguna untuk mengakses kolam pengetahuan mana yang relevan untuk dapat digunakan pada proyek konsultasi selanjutnya. Adapun gambaran sistem yang akan dikembangkan ditunjukkan seperti pada gambar 1.1.



Gambar 1 sistem yang akan dikembangkan

Seperti terlihat pada gambar 1.1, proses dimulai dari format masukan dari pengelola yang menerima data masukan, adapun yang menjadi data masukan pada sistem adalah data dokumen hasil proyek konsultasi dan data pencarian. Selanjutnya data dokumen (data topik konsultasi, data informasi umum perusahaan, dan text isi dokumen) akan diproses untuk disimpan pada media penyimpanan untuk dokumen, dan data dokumen (text isi dokumen) akan diproses melalui proses ekstraksi isi dokumen yang meliputi tahapan casefolding, tokenizing, filtering, dan analyzing, hasil dari proses ini akan tersimpan pada media penyimpanan untuk kata.

Sedangkan untuk data masukan dari pengguna yaitu data pencarian (data topik konsultasi, data informasi umum perusahaan, dan kata kunci) yang akan dimasukkan ke dalam proses pencarian, proses ini memiliki beberapa tahapan, yaitu pertama melakukan pengelompokan data sesuai dengan data topik konsultasi dan data informasi umum perusahaan. Kemudian melakukan tahapan ekstraksi kata kunci untuk kata kunci yang dimasukan, dan tahap terakhir adalah menghitung tingkat kemiripan data hasil seleksi dengan kata kunci yang dimasukan.

Hasil akhir atau output dari sistem yang hendak dicapai, yaitu didapatkannya daftar dokumen yang memiliki tingkat kemiripan yang tinggi antara kata kunci yang dimasukan oleh pengguna sistem (konsultan) dengan dokumen-dokumen yang tersimpan pada sistem, sehingga dapat mempermudah untuk mendapatkan kembali informasi dan pengetahuan yang relevan pada dokumen-dokumen tersebut untuk digunakan ulang pada penanganan proyek konsultasi berikutnya.

3. Analisa dan Hasil

3.1. Proses Text Mining

Proses text mining digunakan untuk mengelompokan kata-kata yang dapat mewakili isi dari dokumen, sehingga dapat dilakukan analisa keterhubungan antar kata pada dokumen. Misalnya, terdapat tiga buah dokumen hasil proyek konsultasi, dokumen pertama (D1) memiliki isi “penurunan omset penjualan batik di Cirebon menurun diakibatkan pangsa pasar batik dimasuki oleh batik buatan Cina yang harganya jauh lebih murah”, dokumen ke-dua (D2) memiliki isi “produk kerajinan tangan tembikar Majalengka dapat meraih pasar luar negeri”, dan dokumen ke-tiga (D3) memiliki isi “diperlukan inovasi rasa pada dodol Garut agar selera pasar tidak hilang dan pasar akan rame kembali”

Setelah dokumen-dokumen tersebut melalui tahapan pre-processing yaitu case folding, tokenizing, filtering, dan analyzing, maka akan didapatkan informasi berupa :

3.1.2. Frekuensi kemunculan kata pada dokumen (tf)

Tabel 3.1 Tabel frekuensi kemunculan kata pada dokumen.

kata	dokumen	tf
menurun	1	1
diakibatkan	1	1
pangsa	1	1
pasar	1	1
dimasuki	1	1
buatan	1	1
Cina	1	1
harganya	1	1
jauh	1	1
murah	1	1
Cirebon	1	1
batik	1	3
penjualan	1	1
omset	1	1
penurunan	1	1
produk	2	1
kerajinan	2	1
tangan	2	1

kata	dokumen	tf
Tembikar	2	1
Majalengka	2	1
Meraih	2	1
Luar	2	1
Negeri	2	1
Pasar	2	1
Pasar	3	2
Diperlukan	3	1
Inovasi	3	1
Rasa	3	1
Dodol	3	1
Garut	3	1
Selera	3	1
Hilang	3	1
Rame	3	1
Kembali	3	1

3.1.3. Frekuensi jumlah dokumen yang mengandung kata yang sama (df)

Tabel 3.2 Tabel frekuensi jumlah dokumen yang mengandung kata yang sama.

kata	df
diperlukan	1
negeri	1
luar	1
meraih	1
majalengka	1
tembikar	1

kata	df
penjualan	1
batik	1
Cirebon	1
menurun	1
diakibatkan	1
pangsa	1

kata	df
Cina	1
harganya	1
jauh	1
murah	1
produk	1
kerajinan	1

kata	df
Garut	1
selera	1
hilang	1
rame	1
kembali	1

tangan	1	pasar	3	inovasi	1
penurunan	1	dimasuki	1	rasa	1
omset	1	buatan	1	dodol	1

3.2. Search Engine (Mesin Pencari)

Proses pencarian pada sistem menggunakan metode TF-IDF (Term Frequency - Inverse Dokumen Frequency) dan Cosine. Melalui perhitungan frekuensi kemunculan tiap kata kunci pada dokumen, sehingga akan didapatkan bobot dari dokumen-dokumen yang memiliki similiaritas dengan kata kunci.

Adapun formulasi yang digunakan untuk menghitung bobot masing-masing dokumen terhadap kata kunci adalah :

$$W_{d,t} = tf_{d,t} * IDF_t \tag{1}$$

Dimana :

- W = bobot dokumen ke-d terhadap kata ke-t
- tf = banyaknya kata yang dicari pada sebuah dokumen
- IDF_t = log (D/df)
- D = total dokumen yang mengandung kata kunci
- df = banyaknya dokumen yang mengandung kata kunci yang dicari
- d = dokumen ke-d
- t = kata ke-t dari kata kunci

Misalnya, terdapat masukan kata kunci “pasar batik”, maka sistem akan melakukan tahapan preprocessing terlebih dahulu terhadap kata kunci tersebut, sehingga hasil akhir akan menjadi token-token kata “pasar” dan “batik”, maka setelah itu bobot dari masing-masing dokumen terhadap kata kunci dapat dihitung dan diakumulasikan, sedangkan untuk menghindari nilai perhitungan IDF yang menghasilkan nol, karena total dokumen yang mengandung kata kunci (D) mempunyai nilai yang sama dengan banyaknya dokumen yang mengandung kata kunci yang dicari (df), maka setiap hasil dari IDF akan ditambah satu. Sehingga formulasi untuk menghitung IDF menjadi :

$$IDF_t = \log(D/df_t) + 1 \tag{2}$$

Dimana :

- t = kata ke-t dari kata kunci
- df = jumlah dokumen kata ke-t dari kata kunci
- D = jumlah dokumen yang mempunyai kata kunci
- IDF = rasio frekuensi dokumen pada kata ke-t dari kata kunci

Hasil yang didapatkan untuk penghitungan masing-masing bobot dokumen terhadap kata kunci dengan menggunakan contoh dokumen pada sub bab 3.4 dapat dilihat pada Tabel 3.3.

Tabel 3.3 Tabel perhitungan TF-IDF

token	tf				df	D/df	IDF	W			
	kk	D1	D2	D3				Kk	D1	D2	D3
diperlukan	0	0	0	1	1	3	1.477	0	0	0	1.477
negeri	0	0	1	0	1	3	1.477	0	0	1.477	0
luar	0	0	1	0	1	3	1.477	0	0	1.477	0
meraih	0	0	1	0	1	3	1.477	0	0	1.477	0
majalengka	0	0	1	0	1	3	1.477	0	0	1.477	0
tembikar	0	0	1	0	1	3	1.477	0	0	1.477	0
tangan	0	0	1	0	1	3	1.477	0	0	1.477	0
penurunan	0	1	0	0	1	3	1.477	0	1.477	0	0
omset	0	1	0	0	1	3	1.477	0	1.477	0	0
penjualan	0	1	0	0	1	3	1.477	0	1.477	0	0

batik	1	3	0	0	1	3	1.477	1.477	4.431	0	0
cirebon	0	1	0	0	1	3	1.477	0	1.477	0	0
menurun	0	1	0	0	1	3	1.477	0	1.477	0	0
diakibatkan	0	1	0	0	1	3	1.477	0	1.477	0	0
pangsa	0	1	0	0	1	3	1.477	0	1.477	0	0
pasar	1	1	1	2	3	1	1	1	1	1	2
dimasuki	0	1	0	0	1	3	1.477	0	1.477	0	0
buatan	0	1	0	0	1	3	1.477	0	1.477	0	0
cina	0	1	0	0	1	3	1.477	0	1.477	0	0
harganya	0	1	0	0	1	3	1.477	0	1.477	0	0
jauh	0	1	0	0	1	3	1.477	0	1.477	0	0
murah	0	1	0	0	1	3	1.477	0	1.477	0	0
produk	0	0	1	0	1	3	1.477	0	0	1.477	0
kerajinan	0	0	1	0	1	3	1.477	0	0	1.477	0
inovasi	0	0	0	1	1	3	1.477	0	0	0	1.477
rasa	0	0	0	1	1	3	1.477	0	0	0	1.477
dodol	0	0	0	1	1	3	1.477	0	0	0	1.477
garut	0	0	0	1	1	3	1.477	0	0	0	1.477
selera	0	0	0	1	1	3	1.477	0	0	0	1.477
hilang	0	0	0	1	1	3	1.477	0	0	0	1.477
rame	0	0	0	1	1	3	1.477	0	0	0	1.477
kembali	0	0	0	1	1	3	1.477	0	0	0	1.477

Maka bobot untuk masing-masing dokumen setelah diakumulasikan adalah :

- D1 = 1+4.431 = 5.431
- D2 = 1+0 = 1
- D3 = 1+0 = 2

Tahap berikutnya agar bobot nilai lebih presisi digunakan penghitungan nilai masing-masing dokumen dengan menggunakan formulasi Cosine :

$$cosine Di = \frac{\sum_{j=1}^n (kkj * Di, j)}{(sqrt(kk) * sqrt(Di))} \tag{3}$$

Tabel 3.4 Tabel perhitungan vector sapce model

token	Kk ²	D1 ²	D2 ²	D3 ²	kk*D1	kk*D2	kk*D3
pasar	1	1	1	4	1	1	2
batik	2,19	19,71	0	0	6.57	0	0
penurunan	0	2.18	0	0	0	0	0
omset	0	2.18	0	0	0	0	0
penjualan	0	2.18	0	0	0	0	0
cirebon	0	2.18	0	0	0	0	0
menurun	0	2.18	0	0	0	0	0
diakibatkan	0	2.18	0	0	0	0	0
pangsa	0	2.18	0	0	0	0	0
pasar	0	2.18	0	0	0	0	0
dimasuki	0	2.18	0	0	0	0	0
buatan	0	2.18	0	0	0	0	0
cina	0	2.18	0	0	0	0	0
harganya	0	2.18	0	0	0	0	0
jauh	0	2.18	0	0	0	0	0
murah	0	2.18	0	0	0	0	0
negeri	0	0	2.18	0	0	0	0
luar	0	0	2.18	0	0	0	0
meraih	0	0	2.18	0	0	0	0

majalengka	0	0	2.18	0	0	0	0
tembikar	0	0	2.18	0	0	0	0
tangan	0	0	2.18	0	0	0	0
produk	0	0	2.18	0	0	0	0
kerajinan	0	0	2.18	0	0	0	0
diperlukan	0	0	0	2.18	0	0	0
inovasi	0	0	0	2.18	0	0	0
rasa	0	0	0	2.18	0	0	0
dodol	0	0	0	2.18	0	0	0
garut	0	0	0	2.18	0	0	0
selera	0	0	0	2.18	0	0	0
hilang	0	0	0	2.18	0	0	0
rame	0	0	0	2.18	0	0	0
kembali	0	0	0	2.18	0	0	0
	Sqrt($\sum K_k$)	Sqrt($\sum D_i^2$)	$\sum(KK * D_i)$				
total	1.78	7.157	4.29	4.86	7.57	1	2

Maka nilai cosine dari tiap-tiap dokumen adalah :

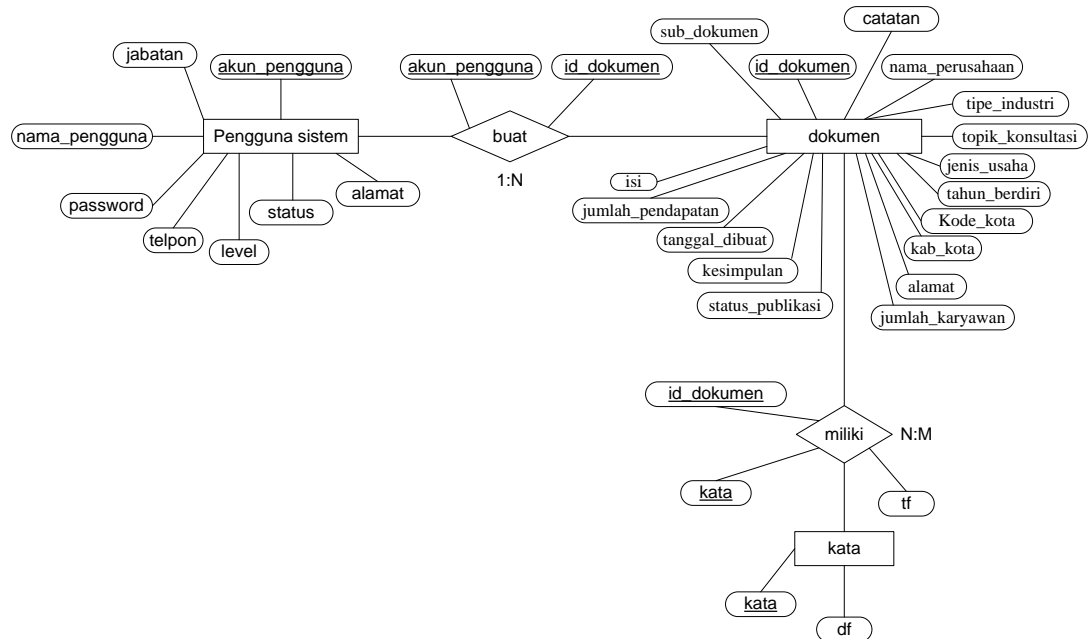
- $D1 = 7.57 / (1.78 * 7.157) = 0.59$
- $D2 = 1 / (1.78 * 4.29) = 0.1309$
- $D3 = 2 / (1.78 * 4.86) = 0.23$

Semakin besar nilai cosine yang dimiliki oleh dokumen, maka akan semakin tinggi juga tingkat similiaritas kata kunci dengan dokumen tersebut.

3.3. Perancangan Basis Data

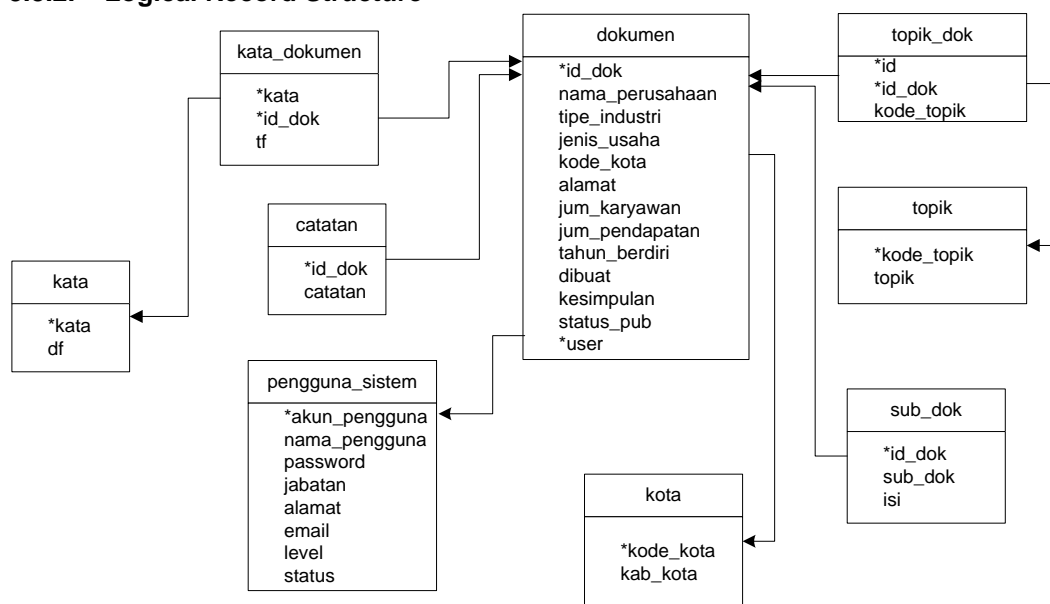
3.3.1. Entity Relationship Diagram (ERD)

Pembentukan ERD pada dilakukan dengan menganalisa entitas yang terlibat pada sistem yang akan dikembangkan. Entity Relationship Diagram untuk KMS - Hasil Proyek Konsultasi dapat dilihat pada Gambar 2.



Gambar 2. ERD KMS - hasil proyek konsultasi

3.3.2. Logical Record Structure



Gambar 3.12 Logical Record Structure Hasil Normalisasi

4. Kesimpulan

Penyimpanan, pengelolaan, dan pengintegrasian data pada dokumen-dokumen hasil proyek konsultasi telah berhasil dilakukan dengan menggunakan algoritma text mining. Proses ekstraksi data (isi dokumen) telah berhasil dilakukan menggunakan algoritma text mining untuk disimpan pada sistem melalui tahapan-tahapan dalam text mining, yaitu case folding, tokenizing, filtering, dan analyzing. sehingga dengan bantuan penggunaan metode TF-IDF(Term Frequency-Inverse Document Frequency) dan Cosine dari mode VSM (Vector Space Model) pada IR(Information Retrieval) telah berhasil memberikan result pencarian dokumen-dokumen yang memiliki tingkat kemiripan dengan kata kunci yang dimasukkan pada sistem. Sehingga, pengguna dapat memperoleh kembali informasi dan pengetahuan yang tersimpan pada dokumen dengan cepat, tepat dan mudah.

Daftar Pustaka

- [1] Alavi, M., Leidner, D.E."Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues". MIS Quarterly. 2001.
- [2] Ariwibowo, Arnold., Lukas, Samuel., Fransiska, Lydia., "Rancang Bangun Piranti Lunak Otomasi Perbandingan Teks", konferensi Nasional sistem dan informatika 2009; Bali, November 14, 2009.
- [3] Hermawansyah, Adit, "Aplikasi Pengkategorian Dokumen Dan Pengukuran Tingkat Similaritas Dokumen Menggunakan Kata Kunci Pada Dokumen Penulisan Ilmiah Universitas Gunadarma", Fakultas Ilmu Komputer, Universitas Gunadarma. 2009
- [4] Fakhur Rozi, Nanang, "Identifikasi Jenis Hadits Menggunakan Beberapa Kombinasi Metode Learning", Jurusan Teknik Informatik PENS-ITS Surabaya. 2010
- [5] Nonaka, Ikujiro., [Takeuchi, Hirotaka.](#) "[The knowledge creating company: how Japanese companies create the dynamics of innovation](#)", New York: Oxford University Press. 1995.
- [6] Awad, ghaziri. "International Edition : Knowledge Management", Prentice Hall, 2001.
- [7] Feldman, Ronan "The Text Mining Hand Book : Advance Approaches in Analyzing Unstructured Data", Hebrew University of Jerusalem Jamer Sanger, ABS Ventress, Boston, Massachusetts. 2007.