

IMPLEMENTASI *TEXT MINING* UNTUK MENGETAHUI *RESEARCH TREND* MAHASISWA JURUSAN SISTEM INFORMASI UNIVERSITAS SRIWIJAYA

SKRIPSI

Sebagai salah satu syarat untuk penyelesaian
Studi di Program Studi Sistem Informasi S1



Oleh

DAHLIA

09031181520117

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
DESEMBER 2019**

LEMBAR PENGESAHAN

SKRIPSI

IMPLEMENTASI *TEXT MINING* UNTUK MENGETAHUI *RESEARCH TREND* MAHASISWA JURUSAN SISTEM INFORMASI UNIVERSITAS SRIWIJAYA

Sebagai salah satu syarat untuk penyelesaian
Studi di Program Studi Sistem Informasi S1

Oleh

DAHLIA

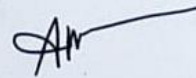
09031181520117

Pembimbing I

Indralaya, Desember 2019
Pembimbing II



Rahmat Izwan Heroza, M.T
NIP. 198706302015041001



Allsela Meiriza, M.T
NIP. 198305132015012201

Mengetahui,
Ketua Jurusan Sistem Informasi



Endang Lestari Ruskan, S.Kom., M.T.
NIP. 197811172006042001

LEMBAR PERSETUJUAN

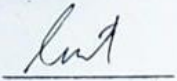
Telah diuji dan lulus pada :

Hari : Jumat

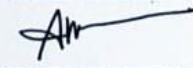
Tanggal : 27 Desember 2019

Tim Penguji

1. Pembimbing I : Rahmat Izwan Heroza, M.T



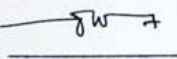
2. Pembimbing II : Allsela Meiriza, M.T



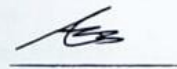
3. Ketua Penguji : Yadi Utama, M.Kom



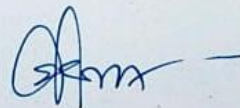
4. Anggota I : Dwi Rosa Indah, M.T



5. Anggota II : Ali Bardadi, M.Kom



**Mengetahui,
Ketua Jurusan Sistem Informasi**



**Endang Lestari Ruskan, M.T
NIP 197811172006042001**

HALAMAN PERSEMBAHAN

MOTTO :

“Ketika sedang mengalami kesusahan dan bertanya-tanya kemana Allah, cukup ingat bahwa seorang guru selalu diam saat ujian berjalan” –Nourman Ali Khan

Skripsi ini kupersembahkan kepada:

- *Orang tua tercinta*
- *Sahabat-sahabatku*
- *Keluarga Besar*
- *Dosen Pembimbing Skripsi*
- *Almamaterku, Universitas Sriwijaya*

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Dahlia
NIM : 09031181520117
Program Studi : Sistem Informasi Reguler
Judul Skripsi : Implementasi *Text Mining* untuk Mnengetahui *Research Trend* Mahasiswa Jurusan Sistem Informasi Universitas Sriwijaya

Hasil Pengecekan *Software iThenticate/Turnitin* : 18 %

Menyatakan bahwa laporan skripsi saya merupakan hasil karya saya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan skripsi ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya dengan ketentuan yang berlaku.

Demikianlah pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Indralaya, Desember 2019
METERAI TEMPEL
16CFADAHF1634926
6000
Dahlia
NIM. 09031181520117

KATA PENGANTAR



Alhamdulillah rabbil'alamini, puji dan syukur kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya serta memberikan kesehatan, kekuatan, dan kesabaran sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul **“IMPLEMENTASI *TEXT MINING* UNTUK MENGETAHUI *RESEARCH TREND* MAHASISWA JURUSAN SISTEM INFORMASI UNIVERSITAS SRIWIJAYA”**.

Dalam melaksanakan dan menyusun Tugas Akhir ini, penulis ingin menyampaikan ucapan terimakasih kepada :

1. Allah SWT yang telah memberikan nikmat kesehatan dengan rasa syukur alhamdulillah sehingga penulis dapat menyelesaikan Tugas Akhir ini.
2. Kedua Orang Tua tercinta, terima kasih atas segala cinta kasih dan sayang yang tiada hentinya yang selalu senantiasa memberikan semangat, dukungan dan doa sehingga penulis dapat menyelesaikan Tugas Akhir ini.
3. Bapak Jaidan Jauhari, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Ibu Endang Lestari Ruskan, M.T. selaku Ketua Jurusan Sistem Informasi.
5. Bapak Rahmat Izwan Heroza, M.T. dan Ibu Allsela Meiriza, M.T. selaku Dosen Pembimbing yang telah membimbing dan memberikan masukan serta ide yang membangun sehingga Tugas Akhir ini dapat diselesaikan.
6. Ibu Dwi Rosa Indah, M.T., Bapak Ali Bardadi M.Kom., Bapak Yadi Utama, M.Kom., dan Ibu Ken Ditha Tania M.Kom., selaku Dosen Penguji yang memberikan kritik serta saran yang membangun sehingga membuat Tugas Akhir ini semakin lebih baik.

7. Seluruh Dosen yang telah mengajarkan, membimbing dan membekali ilmu kepada penulis selama menuntut ilmu di Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.
8. Sahabat – sahabat seperjuangan yang telah mendukung, memberi motivasi serta menjadi pendengar yang baik selama penyusunan Tugas Akhir ini.
9. Seluruh teman – teman Sistem Informasi Reguler 2015 yang memberikan motivasi maupun dorongan untuk menyelesaikan Tugas Akhir ini.

Penulis menyadari bahwa Tugas Akhir ini masih jauh dari kesempurnaan, baik teknis penulisan, bahasa maupun cara pemaparannya. Oleh karena itu kritik dan saran dari semua pihak yang bersifat membangun sangat penulis harapkan demi kesempurnaan karya tulis ini. Akhir kata, penulis mengucapkan terima kasih dan mohon maaf kepada semua pihak yang namanya tak bisa disebutkan satu per satu. Semoga Allah membalas semua kebaikannya.

Indralaya, Desember 2019
Penulis,

Dahlia
NIM. 09031181520117

**IMPLEMENTASI *TEXT MINING* UNTUK MENGETAHUI
RESEARCH TREND MAHASISWA JURUSAN SISTEM
INFORMASI UNIVERSITAS SRIWIJAYA**

Oleh

Dahlia 09031181520117

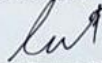
ABSTRAK

Pada bidang studi sistem informasi, terdapat banyak topik yang dijadikan tema penelitian skripsi mahasiswa. Sayangnya, tidak jarang mahasiswa mengambil tema skripsi yang sama dengan tema skripsi mahasiswa angkatan sebelumnya. Karena memiliki kesamaan tema, akhirnya muncullah judul-judul skripsi yang hampir mirip. Hal ini tentu membuat penelitian di kalangan mahasiswa tidak dapat berkembang dengan cepat. Dan juga apakah penelitian ini sudah sesuai dengan *Body of Knowledge* Sistem Informasi yang seharusnya dapat dijadikan acuan dalam melakukan suatu penelitian pada bidang Sistem Informasi agar sesuai dengan disiplin ilmu Sistem Informasi. Penelitian ini bertujuan untuk menganalisis tema-tema skripsi dari teks abstrak yang terdapat pada penelitian mahasiswa jurusan Sistem Informasi Universitas Sriwijaya pada rentang waktu tertentu. Pengolahan teks abstrak dilakukan dengan teknik klasifikasi teks menggunakan algoritma *naive bayes*. Hasil dari penelitian ini adalah penelitian mahasiswa terkelompok berdasarkan kategori *body of knowledge* sistem informasi dan diketahui tren dan perkembangan penelitian dari tahun ke tahun.

Kata Kunci : *text mining*, klasifikasi teks, *naive bayes*, tren penelitian.

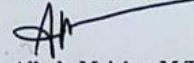
Indralaya, Desember 2019

Pembimbing I,




Rahmat Izwan Heroza, M.T
NIP 198706302015041001

Pembimbing II,



Allsela Meiriza, M.T
NIP 198305132015012201

Mengetahui,
Ketua Jurusan Sistem Informasi,



Endang Lestari Ruskan, M.T
NIP. 197811172006042001

**IMPLEMENTATION OF TEXT MINING TO FIND OUT
RESEARCH TRENDS OF STUDENTS MAJORING IN
INFORMATION SYSTEMS AT SRIWIJAYA UNIVERSITY**

By

Dahlia 09031181520117

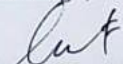
ABSTRACT

In the field of information systems studies, there are many topics that are used as research themes for student thesis. Unfortunately, it is not uncommon for students to take the same thesis theme as the previous generation student's thesis theme. Because it has a common theme, finally came the thesis titles that are almost similar. This certainly makes research among students unable to develop quickly. And also whether this research is in accordance with the Body of Knowledge Information Systems that should be used as a reference in conducting a research in the field of Information Systems to fit the discipline of Information Systems. This study aims to analyze the thesis themes of abstract texts contained in the research of students majoring in Information Systems at the University of Sriwijaya in a certain time span. Abstract text processing is done by text classification techniques using the naïve Bayes algorithm. The results of this study are student studies grouped according to the information system body of knowledge category and known research trends and developments from year to year.

Keywords : text mining, text classification, naïve bayes, research trend.

Indralaya, Desember 2019

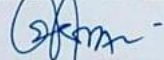
Pembimbing I,


Rahmat Izwan Heroza, M.T
NIP 198706302015041001

Pembimbing II,


Allsela Meiriza, M.T
NIP 198305132015012201

Mengetahui,
Ketua Jurusan Sistem Informasi,


Endang Lestari Ruskan, M.T
NIP. 197811172006042001

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PERSEMBAHAN	iv
HALAMAN PERNYATAAN	v
KATA PENGANTAR	vi
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan.....	3
1.4 Manfaat.....	3
1.5 Batasan Masalah.....	4
BAB II TINJAUAN PUSTAKA	5
2.1 <i>Data Mining</i>	5
2.1.1 Pengertian <i>Data Mining</i>	5
2.1.2 Arsitektur <i>Data Mining</i> (Penambangan Data).....	5
2.1.3 Metodologi <i>Data Mining</i> (Penambangan Data).....	6
2.2 <i>Text Mining</i>	9
2.3 Klasifikasi.....	9
2.3.1 Model klasifikasi.....	10

2.4 Naïve Bayes Classifier	10
2.5 Skripsi.....	11
2.6 <i>Body of Knowledge</i> (BOK).....	12
BAB III METODELOGI PENELITIAN.....	16
3.1 Kerangka Pikir penelitian	16
3.2 Teknik Pengumpulan Data	17
3.2.1 Jenis Data	17
3.2.2 Sumber Data.....	17
3.2.3 Metode Pengumpulan Data.....	17
3.3 Metode Penelitian.....	18
BAB IV ANALIS SISTEM	20
4.1 Fase Pemahaman Bisnis	20
4.1.1 Kebutuhan Fungsional	20
4.1.2 Kebutuhan Non Fungsional	20
4.1.3 Analisa Proses Bisnis.....	21
4.1.3.1.1 Diagram Dekomposisi.....	21
4.1.3.1.2 Diagram Konteks	22
4.1.3.1.3 DFD Level 1.....	23
4.1.3.1.4 DFD Level 2 Mengelola Data <i>User</i>	24
4.1.3.1.5 DFD Level 2 Mengelola Data Tugas Akhir.....	25
4.2 Fase Pemahaman Data.....	26
4.2 Persiapan Data.....	26
4.2.1 Entity Relational Diagram (ERD).....	27
4.2.2 Rancangan Desain Tabel Database.....	28
4.5 Fase Pemodelan	31
BAB V PERANCANGAN SISTEM	41
5.1 Physical Data Flow Diagram (<i>PDFD</i>)	41
5.2 Database Design	43
5.3 Rancangan <i>Interface</i>	50
5.3.1 Halaman Login.....	50
5.3.2 Halaman Beranda Admin.....	51

5.3.3	Halaman Kelola Akun Admin.....	51
5.3.4	Halaman Tambah Akun Admin	52
5.3.5	Halaman Kelola Tugas Akhir Admin	52
5.3.6	Halaman Tambah Data Tugas Akhir Admin	53
5.3.7	Halaman Edit Data Tugas Akhir Admin.....	54
5.3.8	Halaman Tren Admin.....	55
5.3.9	Halaman Daftar Akun	56
5.3.10	Halaman Beranda Mahasiswa	56
5.3.11	Halaman Kelola Tugas Akhir Mahasiswa.....	57
5.3.12	Halaman Tambah Data Tugas Akhir Mahasiswa	58
5.3.13	Halaman Profil Mahasiswa	59
5.3.14	Halaman Edit Data Diri Mahasiswa.....	59
5.3.15	Halaman Awal.....	60
5.3.16	Halaman Cari Tugas Akhir	60
BAB VI HASIL DAN PEMBAHASAN		61
6.1	Hasil.....	61
6.2	Pembahasan	61
6.2.1	Halaman Login.....	61
6.2.2	Halaman Beranda Admin.....	62
6.2.3	Halaman Kelola Akun Admin.....	63
6.2.4	Halaman Tambah Akun Admin	64
6.2.5	Halaman Kelola Tugas Akhir Admin	64
6.2.6	Halaman Tambah Data Tugas Akhir Admin	65
6.2.7	Halaman Edit Data Tugas Akhir Admin.....	66
6.2.8	Halaman Tren Admin.....	67
6.2.9	Halaman Daftar Akun	68
6.2.10	Halaman Beranda Mahasiswa	68
6.2.11	Halaman Kelola Tugas Akhir Mahasiswa.....	69
6.2.12	Halaman Tambah Data Tugas Akhir Mahasiswa	70
6.2.13	Halaman Profil Mahasiswa	71
6.2.14	Halaman Edit Data Diri Mahasiswa.....	71

6.2.15	Halaman Awal.....	72
6.2.16	Halaman Cari Tugas Akhir	72
6.3	Pengujian Sistem	73
BAB VII KESIMPULAN DAN SARAN.....		76
7.1	Kesimpulan.....	76
7.2	Saran.....	77
DAFTAR PUSTAKA		78

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Siklus CRISP-DM	6
Gambar 2.2 Proses Klasifikasi dengan Naïve Bayes	10
Gambar 3.1 Kerangka pikir penelitian	16
Gambar 4.1 Diagram Dekomposisi	22
Gambar 4.2 Diagram Konteks	22
Gambar 4.3 DFD Level 1	23
Gambar 4.4 DFD Level 2 Mengelola Data User	24
Gambar 4.5 DFD Level 2 Mengelola Data Tugas Akhir	25
Gambar 4.6 <i>Entity Relational Diagram</i>	27
Gambar 5.1 PDFD Level 2 Mengelola Data User	41
Gambar 5.2 PDFD Level 2 Mengelola Data Tugas Akhir	42
Gambar 5.3 Skema <i>Database</i>	43
Gambar 5.4 Halaman <i>Login</i>	50
Gambar 5.4 Halaman Beranda Admin	51
Gambar 5.5 Halaman Kelola Akun Admin	51
Gambar 5.6 Halaman Tambah Akun Admin	52
Gambar 5.7 Halaman Kelola TA Admin	52
Gambar 5.8 Halaman Tambah Data TA Admin	53
Gambar 5.9 Halaman Edit Data TA Admin	54
Gambar 5.10 Halaman Tren Admin	55
Gambar 5.11 Halaman Daftar Akun.....	56
Gambar 5.12 Halaman Daftar Akun	56
Gambar 5.13 Halaman Kelola TA Mahasiswa	57
Gambar 5.14 Halaman Tambah TA Mahasiswa	58
Gambar 5.15 Halaman Profil	59
Gambar 5.16 Halaman Edit Data Diri Mahasiswa	59

Gambar 5.17 Halaman Awal	60
Gambar 5.17 Cari TA.....	60
Gambar 6.1 Halaman Login	61
Gambar 6.2 Halaman Beranda Admin	62
Gambar 6.3 Halaman Kelola Akun Admin	63
Gambar 6.4 Halaman Tambah Akun Admin	64
Gambar 6.5 Halaman Kelola TA Admin	64
Gambar 6.6 Halaman Tambah Data TA Admin	65
Gambar 6.7 Halaman Edit Data TA Admin	66
Gambar 6.8 Halaman Tren Admin	67
Gambar 6.9 Halaman Daftar Akun	68
Gambar 6.10 Halaman Beranda Mahasiswa	68
Gambar 6.11 Halaman Kelola TA Mahasiswa	69
Gambar 6.12 Halaman Tambah Data TA Mahasiswa	70
Gambar 6.13 Halaman Profil Mahasiswa	71
Gambar 6.14 Halaman Edit Data Diri Mahasiswa	71
Gambar 6.15 Halaman Awal.....	72
Gambar 6.15 Halaman Cari TA.....	72

DAFTAR TABEL

	Halaman
Tabel 2.1 <i>Body of Knowledge</i> Sistem Informasi	12
Tabel 2.2 Penjabaran Aspek Pengetahuan spesifik sistem informasi	13
Tabel 4.1 Kebutuhan Fungsional	20
Tabel 4.2 Kebutuhan Non-fungsional	20
Tabel 4.3 Atribut pada raw data	26
Tabel 4.4 <i>User</i>	28
Tabel 4.5 Topik	29
Tabel 4.6 Tugas Akhir	30
Tabel 4.7 Contoh Isi <i>Stopword</i>	31
Tabel 4.8 Contoh Kata Dasar <i>Stemming</i>	33
Tabel 4.9 Tabel Teks Abstrak	34
Tabel 6.1 Pengujian <i>black box</i>	73

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Surat Keputusan Tugas Akhir.....	A-1
Lampiran 2 Kartu Konsultasi Tugas Akhir.....	B-2

BAB I

PENDAHULUAN

1.1 Latar Belakang

Menurut Ronald Stamper, Sistem Informasi adalah bidang ilmu yang menjembatani kesenjangan antara komponen teknis (TI) dan non teknis (manajemen). Sistem Informasi memiliki cakupan bidang ilmu yang telah dikaji oleh ACM dan AIS dan menerbitkan *Body of Knowledge* (BOK) Sistem Informasi sebagai basis sebuah Disiplin Ilmu sekaligus sebagai pengetahuan bagi profesi Sistem Informasi.

Universitas Sriwijaya merupakan salah satu perguruan tinggi yang memiliki Jurusan Sistem Informasi. Setiap tahunnya Jurusan Sistem Informasi Universitas Sriwijaya menghasilkan lulusan-lulusan Sistem Informasi yang telah melakukan suatu penelitian yang disebut skripsi. Pada bidang studi sistem informasi, terdapat banyak topik yang dijadikan tema penelitian skripsi mahasiswa. Namun apakah penelitian ini sudah sesuai dengan BOK Sistem Informasi. BOK Sistem Informasi seharusnya dapat dijadikan acuan dalam melakukan suatu penelitian pada bidang Sistem Informasi agar sesuai dengan disiplin ilmu Sistem Informasi.

Pada setiap karya ilmiah terdapat abstrak yang merupakan sintesa keseluruhan dari isi dokumen karya ilmiah (Nasution, 2017). Abstrak adalah bagian yang bisa digunakan pembaca untuk mengetahui isi dasar dokumen dengan cepat dan akurat tanpa harus membaca seluruh isi dokumen. Dengan mengidentifikasi bagian abstrak sudah bisa di dapat tema dan isi dari penelitian yang dilakukan oleh penulis. Dengan

adanya teks abstrak ini dapat dimanfaatkan untuk diolah menggunakan teknik *data mining*.

Text mining merupakan cabang dari bidang *data mining* dan pada *text mining* data yang digunakan berupa teks (Mustikasari, 2017). *Text mining* mengacu pada mengacu pada metode menemukan informasi yang bermakna dengan menganalisis data teks (Perveen *et al.*, 2016). *Text mining* bertujuan mendapatkan informasi baru dari teks tertulis. *Text mining* diperlukan untuk mengelola data teks abstrak untuk mendapatkan mengelompokkan tema sehingga tema skripsi dapat dipetakan. Salah satu metode pengelompokkan *text mining* yaitu klasifikasi. Pada klasifikasi, dokumen akan dikelompokkan berdasarkan kemiripannya atau kesamaannya, kategori kelompok ditentukan terlebih dahulu. Cara ini dianggap lebih mudah untuk kumpulan data yang diketahui atau bisa diperkirakan kategorinya.

Berdasarkan uraian diatas, penulis tertarik melakukan penelitian Tugas Akhir yang berjudul **“Implementasi *Text Mining* untuk Mengetahui *Research Trend* Mahasiswa Jurusan Sistem Informasi Universitas Sriwijaya”**

1.2 Rumusan Masalah

- a. Bagaimana mengelompokkan data teks abstrak skripsi mahasiswa sistem informasi Universitas Sriwijaya?
- b. Bagaimana tren penelitian yang berkembang di Jurusan Sistem Informasi Universitas Sriwijaya?

1.3 Tujuan

Tujuan dari Penelitian ini adalah untuk mengelompokkan data teks abstrak skripsi mahasiswa Sistem Informasi di Universitas Sriwijaya dengan *text mining* untuk memetakan tren penelitian pada bidang sistem informasi.

1.4 Manfaat

Manfaat yang diharapkan dalam penelitian Tugas Akhir ini, antara lain:

1. Mengetahui tren dan perkembangan penelitian pada bidang sistem informasi Universitas Sriwijaya.
2. Membantu untuk menemukan tema penelitian bagi mahasiswa berdasarkan perkembangan penelitian sistem informasi.
3. Mengetahui Ketercapaian BOK (*Body of Knowledge*) Sistem Informasi pada Jurusan SI Universitas Sriwijaya.
4. Sebagai petunjuk untuk menyusun rencana penelitian jangka panjang Universitas
5. Sebagai bahan kajian untuk pengembangan kurikulum.

1.5 Batasan Masalah

Penulis membatasi ruang lingkup permasalahan penulisan Tugas Akhir ini yaitu:

1. Data yang digunakan adalah teks abstrak skripsi mahasiswa Sistem Informasi Universitas Sriwijaya strata satu beberapa tahun terakhir .
2. Penelitian menggunakan CRISP-DM (Cross industry Standard Process – Data Mining).
3. Penelitian ini mengelompokkan data teks dengan teknik klasifikasi

DAFTAR PUSTAKA

- Aznar-sánchez, J. A. *et al.* (2019) 'The worldwide research trends on water ecosystem services', *Ecological Indicators*. Elsevier, 99(November 2018), pp. 310–323. doi: 10.1016/j.ecolind.2018.12.045.
- Burcu, H. and Alpay, N. (2018) 'Central Bank Review A text mining application on monthly price developments reports', *Central Bank Review*. Elsevier Ltd, 18(2), pp. 51–60. doi: 10.1016/j.cbrev.2018.05.001.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. Waltham: Elsevier.
- Musa, O. (2015) 'Sistem Informasi Pemetaan Pendidikan Menggunakan Algoritma Data Mining', 1, pp. 26–32.
- Mustikasari, D. (2017) 'Analisis Tema Skripsi Mahasiswa Menggunakan Document Clustering Dengan Algoritma LINGO', 2(2), pp. 131–140.
- Nasution, M. K. M. (2017) 'Abstrak - Suatu Karya Ilmiah', (December), pp. 13–16. doi: 10.13140/RG.2.2.14333.26086.
- Nugraha, P. D. and Faraby, S. Al (2018) 'Klasifikasi Dokumen Menggunakan Metode k -Nearest Neighbor (kNN) dengan Information Gain Document Classification using k- Nearest Neighbor (k NN) Method with Information Gain', 5(1), pp. 1541–1550.
- P. Seminar and N. Aplikasi, "Klasifikasi teks dengan naïve bayes classifier (nbc) untuk pengelompokan teks berita dan abstract akademis," no. 2011, pp. 269–277, 2012.
- Perveen, S. *et al.* (2016) 'Performance Analysis of Data Mining Classification Techniques to Predict Diabetes', *Procedia - Procedia Computer Science*. Elsevier Masson SAS, 82(March), pp. 115–121. doi: 10.1016/j.procs.2016.04.016.
- Ruiz-saenz, J. and Martinez-gutierrez, M. (2015) 'Virology research in a Latin American developing country : A bibliometric analysis of virology in Colombia (2000 – 2013)
- Original Article Virology research in a Latin American developing country : a bibliometric analysis of virology in Colombia (2000 – 2013)', (December). doi: 10.3855/jidc.6627.
- Sung, K. and Juyoung, K. (2018) 'Analyzing the discriminative attributes of products using text mining focused on cosmetic reviews', 54(May), pp. 938–957. doi: 10.1016/j.ipm.2018.06.003.