

KLASIFIKASI DOKUMEN SCIENTIFIC PAPER GOOGLE  
SCHOLAR DOSEN SISTEM INFORMASI DENGAN METODE K-  
NEAREST NEIGHBOR

TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat  
untuk Mem peroleh Gelar Sarjana Komputer pada  
Program Studi Sistem Informasi

Oleh:

YOLLA DAMARA

11653200150



UIN SUSKA RIAU

FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS ISLAM  
NEGERI SULTAN SYARIF KASIM RIAU PEKANBARU

2022

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim 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.



UIN SUSKA RIAU

## LEMBAR PERSETUJUAN

### KLASIFIKASI DOKUMEN *SCIENTIFIC PAPER GOOGLE SCHOLAR* DOSEN SISTEM INFORMASI DENGAN METODE *K-NEAREST NEIGHBOR*

#### TUGAS AKHIR

Oleh:

**YOLLA DAMARA**

**11653200150**

Telah diperiksa dan disetujui sebagai laporan Tugas Akhir  
di Pekanbaru, pada tanggal 1 Juli 2022

**Ketua Program Studi**



**Eki Saputra, S.Kom., M.Kom.**

**NIP. 198307162011011008**

**Pembimbing**



**Mustakin, ST., M.Kom.**

**NIK. 130511023**



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 milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

## LEMBAR PENGESAHAN

### KLASIFIKASI DOKUMEN *SCIENTIFIC PAPER GOOGLE* *SCHOLAR* DOSEN SISTEM INFORMASI DENGAN METODE *K-NEAREST NEIGHBOR*

#### TUGAS AKHIR

Oleh:

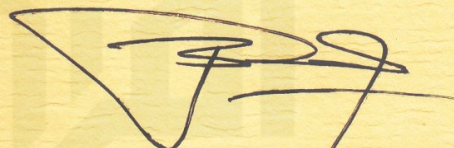
**YOLLA DAMARA**  
**11653200150**

Telah dipertahankan di depan sidang dewan penguji  
sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer  
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau  
di Pekanbaru, pada tanggal 27 Mei 2022

Pekanbaru, 27 Mei 2022

Mengesahkan,

Ketua Program Studi



**Eki Saputra, S.Kom., M.Kom.**  
**NIP. 198307162011011008**



**Dr. Hartono, M.Pd.**  
**NIP. 196403011992030103**

#### DEWAN PENGUJI:

**Ketua : Siti Monalisa, ST., M.Kom.**

**Sekretaris : Mustakim, ST., M.Kom.**

**Anggota 1 : Inggih Permana, ST., M.Kom.**

**Anggota 2 : M. Afdal, ST., M.Kom.**



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 milik UIN Suska Riau

mic University of Sultan Syarif Kasim Riau

## LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum, dengan ketentuan bahwa hak cipta ada pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan atas izin penulis dan harus dilakukan mengikuti kaedah dan kebiasaan ilmiah serta menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin tertulis dari Dekan Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan dapat meminjamkan Tugas Akhir ini untuk anggotanya dengan mengisi nama, tanda peminjaman dan tanggal pinjam pada *form* peminjaman.

### 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.

Lampiran Surat :

Nomor : Nomor 25/2022

Tanggal : 13 Juli 2022

## SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama : Yolla Damara  
NIM : 11653200150  
Tempat / Tanggal Lahir : Sicincin, 31 Maret 1998  
Bidang Studi : Sains dan Teknologi  
Program Studi : Sistem Informasi  
Judul Skripsi : Klasifikasi Dokumen *Scientific Paper* Google Scholar Dosen Sistem Informasi dengan Metode *K-Nearest Neighbor*

Menyatakan dengan sebenar-benarnya bahwa :

1. Penulisan Skripsi dengan judul sebagaimana tersebut diatas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya ilmiah saya sudah disebutkan sumbernya.
3. Oleh karna itu Skripsi saya ini, saya nyatakan bebas dari plagiat.
4. Apabila dikemudian hari terbukti terdapat plagiat dalam penulisan Skripsi saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undang.

Demikian surat pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 13 Juli 2022

Yang membuat pernyataan,



Yolla Damara  
11653200150



© Hak cipta milik 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.

## LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

*Alhamdulillah Rabbil Alamin*, saya ucapkan rasa syukur terhadap Allah SWT yang telah memberikan rahmat, hidayah, ilmu pengetahuan, kesehatan dan kesempatan dalam penyelesaian Tugas Akhir ini. Tugas Akhir ini saya persembahkan kepada:



Segala perjuangan saya hingga saat ini saya hadiahkan kepada orang yang paling berharga dalam hidup saya, yaitu kedua orang tua saya mama dan papa. Saya akan selalu berusaha untuk memberikan yang terbaik atas setiap kepercayaan yang telah mereka berikan. InshaAllah saya akan tumbuh untuk menjadi yang terbaik yang saya bisa. Pencapaian ini menjadi salah satu hadiah kecil dari saya untuk kedua orang tua. Terima kasih atas semua cinta yang diberikan kepada saya. Terima kasih telah menjadi keluarga yang sempurna untuk saya.



Semoga Allah SWT selalu menjaga mereka dimanapun berada, memberikan berkah-Nya untuk setiap langkah, berikan kesehatan dan rezki yang berlimpah kepadanya. Selalu menjadikan mereka hamba yang taat dalam keadaan sesulit apapun. *Aamiin ya Rabbal Aalamiin.*

**YOLLA DAMARA**

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.

## KATA PENGANTAR



*Assalamu'alaikum Warahmatullahi Wabarakatuh. Alhamdulillah, puji dan syukur kehadirat Allah SWT atas berkah, rahmat dan hidyah-Nya yang senantiasa dilimpahkan kepada penulis, sehingga bisa menyelesaikan Tugas Akhir dengan judul "Klasifikasi Dokumen Scientific Paper Google Scholar Dosen Sistem Informasi dengan Metode K-Nearest Neighbor" sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.*

Penyusunan penelitian ini dapat selesai dengan lancar karena tidak lepas dari bantuan berbagai pihak. Untuk itu saya ucapkan terimakasih sebesar-besarnya kepada:

1. Bapak Prof. Dr. Hairunas, M.Ag sebagai Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd sebagai Dekan Fakultas Sains dan Teknologi.
3. Bapak Eki Saputra, S.Kom., M.Kom sebagai Ketua Program Studi Sistem Informasi.
4. Ibu Siti Monalisa, ST., M.Kom selaku Sekretaris Prodi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau dan sebagai Ketua Sidang Tugas Akhir.
5. Bapak Tengku Khairil Ahsyar, S.Kom., M.Kom selaku Kepala Laboratorium Program Studi Sistem Informasi Fakultas Sains dan Teknologi.
6. Bapak Mustakim, ST., M.Kom sebagai Dosen Pembimbing Tugas Akhir ini.
7. Megawati, S.Kom., MT sebagai Pembimbing Akademik yang telah membimbing saya dari awal hingga Semester 4 perkuliahan.
8. M. Afdal, ST., M.Kom sebagai Pembimbing Akademik yang telah membimbing saya dari Semester 5 hingga akhir masa perkuliahan, dan sebagai Dosen Penguji II (dua) yang telah memberi arahan dan kontribusi dalam penyelesaian Tugas Akhir ini.
9. Bapak Inggih Permana, ST., M.Kom sebagai Dosen Penguji I (satu) yang telah memberi arahan dan kontribusi dalam penyelesaian Tugas Akhir ini.
10. Orang tua dan saudara saya yang selalu memberikan doa dan dukungan selama proses pelaksanaan Tugas Akhir ini.



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.

11. Segenap Dosen dan Karyawan Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
12. Semua kawan-kawan yang membantu penulis selama kuliah hingga pelaksanaan Tugas Akhir ini.

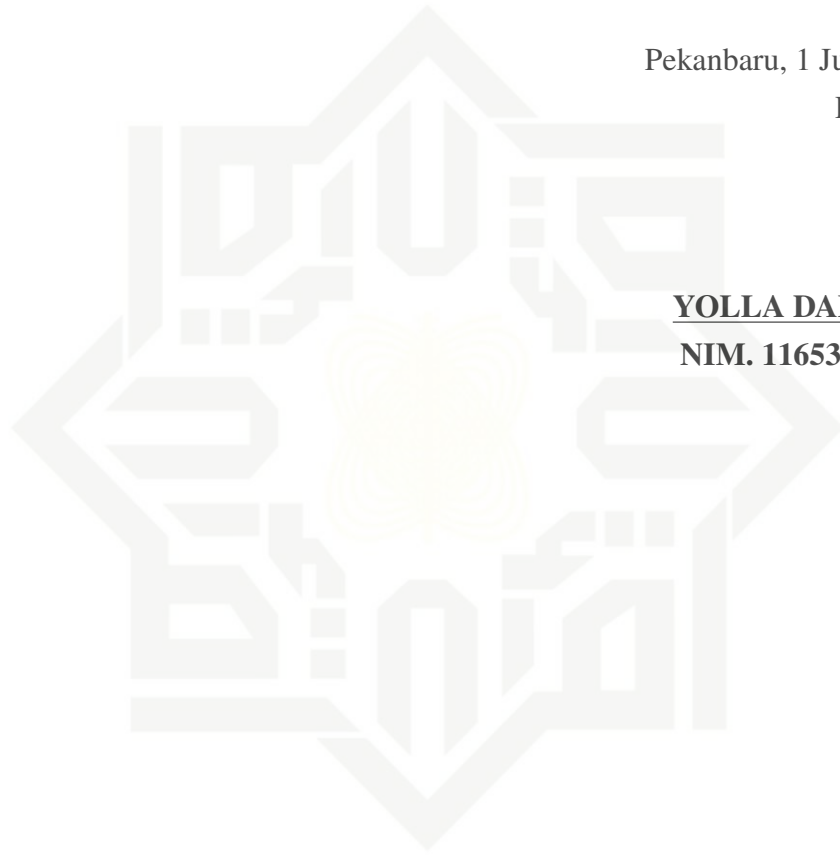
Peneliti tentunya menyadari bahwa pelaksanaan Tugas Akhir ini masih banyak kekurangan dan kelemahannya. Oleh karena itu peneliti berharap kepada semua pihak agar dapat menyampaikan kritik dan saran yang membangun untuk menambah kesempurnaan Tugas Akhir ini. Namun peneliti tetap berharap Tugas Akhir ini dapat bermanfaat bagi semua pihak yang membacanya.

Pekanbaru, 1 Juli 2022

Penulis,

**YOLLA DAMARA**

**NIM. 11653200150**



UIN SUSKA RIAU





UIN SUSKA RIAU

# KLASIFIKASI DOKUMEN *SCIENTIFIC PAPER GOOGLE SCHOLAR* DOSEN SISTEM INFORMASI DENGAN METODE *K-NEAREST NEIGHBOR*

**YOLLA DAMARA**  
**NIM: 11653200150**

Tanggal Sidang: 27 Mei 2022  
Periode Wisuda:

Program Studi Sistem Informasi  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Sultan Syarif Kasim Riau  
Jl. Soebrantas, No. 155, Pekanbaru

## ABSTRAK

Karya tulis ilmiah ialah karya tulis yang telah diakui dalam bidang ilmu pengetahuan, teknologi atau seni yang ditulis atau dikerjakan sesuai dengan tata cara ilmiah dengan mengikuti pedoman dan konvensi ilmiah yang telah disepakati. Dalam karya ilmiah terdapat suatu bidang ilmu pengetahuan yang ditulis untuk menyajikan penelitian. *Scientific paper google scholar* dosen Sistem Informasi UIN SUSKA RIAU dapat digunakan untuk menentukan klasifikasi judul *paper* dosen Sistem Informasi berdasarkan bidang ilmu penelitian dosen dengan menggunakan metode *K-Nearest Neighbor* (K-NN). Tujuan penelitian ini adalah mengklasifikasikan dokumen *scientific paper google scholar* dosen Sistem Informasi sebagai penyetaraan penelitian dosen dengan bidang ilmunya menggunakan algoritma K-NN dengan menggunakan bahasa pemrograman *Python*. Pada penelitian ini dilakukan dua percobaan yaitu menggunakan 3 kelas bidang keahlian, dan menggunakan 22 kelas bidang ilmu. Percobaan tersebut dihitung dengan 253 *data training* dan 109 *data testing*. Berdasarkan percobaan tersebut, kesimpulan yang didapat dari hasil pengklasifikasian *paper Google Scholar* dosen Sistem Informasi yaitu 70.64% *paper* sesuai dengan bidang keahlian serta bidang ilmunya. Pada tingkat akurasi terbaik dengan nilai  $K=10$  untuk kelas bidang ilmu, dan nilai  $k=15$  pada kelas bidang keahlian.

**Kata Kunci:** *Scientific Paper, K-Nearest Neighbor, Google scholar, Python.*

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim 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.



UIN SUSKA RIAU

© Hak cipta milik 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.

# **DOCUMENT CLASSIFICATION OF SCIENTIFIC PAPER GOOGLE SCHOLAR LECTURER OF INFORMATION SYSTEMS USING K-NEAREST NEIGHBOR METHOD**

**YOLLA DAMARA  
NIM: 11653200150**

*Date of Final Exam: Mei 27<sup>th</sup> 2022  
Graduation Period:*

*Department of Information System  
Faculty of Science and Technology  
State Islamic University of Sultan Syarif Kasim Riau  
Soebrantas Street, No. 155, Pekanbaru*

## **ABSTRACT**

*Scientific papers is a written work that has been recognized in the field of science, technology or art written or carried out in accordance with scientific procedures by following agreed scientific guidelines and conventions. In scientific work there is a field of science that is written to present research. Scientific paper google scholar lecturers of information systems UIN SUSKA RIAU can be used to determine the classification of titles of information systems lecturers' papers based on the field of research of lecturers by using the KNN method. The purpose of this research is to classify scientific paper documents, google scholar lecturers, information systems as an equalization of lecturers' research with their fields of knowledge using the K-NN algorithm. Two experiments were carried out in this study, namely: (1) using 3 classes in the field of expertise, and (2) using 22 classes in the field of knowledge, from these experiments will be calculated using the K-Nearest Neighbor (KNN) method. With 253 training data and 109 testing data in each experiment. Based on these experiments, those who have the highest accuracy using the KNN method are 70.64% with a value of  $k = 15$  in 3 classes of expertise and  $k = 10$  in 22 classes of science. For the classification process using the Python programming language.*

**Keywords:** *Scientific Paper, K-Nearest Neighbor, Google scholar, Python.*

## DAFTAR ISI

<b>LEMBAR PERSETUJUAN</b>	<b>ii</b>
<b>LEMBAR PENGESAHAN</b>	<b>iii</b>
<b>LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL</b>	<b>iv</b>
<b>LEMBAR PERNYATAAN</b>	<b>v</b>
<b>LEMBAR PERSEMBAHAN</b>	<b>vi</b>
<b>KATA PENGANTAR</b>	<b>vii</b>
<b>ABSTRAK</b>	<b>ix</b>
<b>ABSTRACT</b>	<b>x</b>
<b>DAFTAR ISI</b>	<b>xi</b>
<b>DAFTAR GAMBAR</b>	<b>xiv</b>
<b>DAFTAR TABEL</b>	<b>xv</b>
<b>DAFTAR SINGKATAN</b>	<b>xvi</b>
<b>1 PENDAHULUAN</b>	<b>1</b>
1.1 Latar Belakang . . . . .	1
1.2 Perumusan Masalah . . . . .	4
1.3 Batasan Masalah . . . . .	4
1.4 Tujuan . . . . .	5
1.5 Manfaat . . . . .	5
1.6 Sistematika Penulisan . . . . .	5
<b>2 LANDASAN TEORI</b>	<b>6</b>
2.1 <i>Text Mining</i> . . . . .	6
2.2 <i>Text Preprocessing</i> . . . . .	6
2.3 <i>Term Frequency–Inverse Document Frequency (TF-IDF)</i> . . . . .	7
2.4 Klasifikasi . . . . .	8
2.5 <i>K-Nearest Neighbor</i> . . . . .	8
2.6 <i>Confusion Matrix</i> . . . . .	9

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim 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.



UIN SUSKA RIAU

2.7	Karya ilmiah . . . . .	9
2.8	<i>Google Scholar</i> atau Cendekia . . . . .	10
2.9	Penelitian Terdahulu . . . . .	10
<b>3</b>	<b>METODOLOGI PENELITIAN</b>	<b>12</b>
3.1	Proses Alur Penelitian . . . . .	12
3.2	Tahapan Perencanaan . . . . .	12
3.3	Proses Pengumpulan Data . . . . .	13
3.4	Tahap Analisa dan Hasil . . . . .	14
3.5	Tahap Dokumentasi . . . . .	14
<b>4</b>	<b>ANALISIS DAN HASIL</b>	<b>15</b>
4.1	Analisa Kebutuhan Data . . . . .	15
4.1.1	Pengumpulan Data . . . . .	15
4.1.2	Data Awal . . . . .	15
4.2	<i>Text Preprocessing</i> . . . . .	15
4.3	<i>Data Training</i> dan <i>Testing</i> . . . . .	19
4.4	Perhitungan TF-IDF . . . . .	20
4.5	Perhitungan <i>K-Nearest Neighbor</i> . . . . .	20
4.5.1	Hasil Akurasi Kelas Bidang Keahlian dengan Nilai K= 5 . . . . .	21
4.5.2	Hasil Akurasi Kelas Bidang Keahlian dengan Nilai K= 10 . . . . .	21
4.5.3	Hasil Akurasi Kelas Bidang Keahlian dengan Nilai K= 15 . . . . .	22
4.5.4	Hasil Akurasi Kelas Bidang Keahlian dengan Nilai K= 20 . . . . .	22
4.5.5	Hasil Akurasi Kelas Bidang Keahlian dengan Nilai K= 25 . . . . .	23
4.5.6	Hasil Akurasi Kelas Bidang Ilmu dengan Nilai K= 5 . . . . .	23
4.5.7	Hasil Akurasi Kelas Bidang Ilmu dengan Nilai K= 10 . . . . .	24
4.5.8	Hasil Akurasi Kelas Bidang Ilmu dengan Nilai K= 15 . . . . .	25
4.5.9	Hasil Akurasi Kelas Bidang Ilmu dengan Nilai K= 20 . . . . .	25
4.5.10	Hasil Akurasi Kelas Bidang Ilmu dengan Nilai K= 25 . . . . .	26
4.6	Analisis Klasifikasi . . . . .	26
<b>5</b>	<b>PENUTUP</b>	<b>29</b>
5.1	Kesimpulan . . . . .	29
5.2	Saran . . . . .	29

**DAFTAR PUSTAKA**

**LAMPIRAN A DAFTAR KATA STOPWORD**

**A - 1**



© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim 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.



## DAFTAR GAMBAR

3.1	Metodologi Penelitian . . . . .	12
4.1	Perbandingan Nilai K pada 3 Kelas . . . . .	27
4.2	Perbandingan Nilai K pada 22 Kelas . . . . .	28
4.3	Perbandingan Nilai K 3 Kelas dan 22 Kelas . . . . .	28



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.



## DAFTAR TABEL

2.1	Tabel Confusion Matrix 2 Kelas . . . . .	9
4.1	Data Awal Dokumen <i>Sciencetific Paper Google Scholar</i> Dosen Sistem Informasi . . . . .	16
4.2	<i>CaseFolding Data Scientific Paper</i> . . . . .	17
4.3	Daftar Kata <i>Stopwords</i> . . . . .	17
4.4	<i>Filtering Data Scientific Paper</i> . . . . .	18
4.5	<i>Preprocessing Data Scientific Paper</i> . . . . .	18
4.6	Data <i>Training</i> kelas Bidang Keahlian menggunakan 3 Kelas . . . . .	19
4.7	Data <i>Testing</i> kelas Bidang Keahlian menggunakan 3 Kelas . . . . .	19
4.8	Pembagian Data Latih dan Data Uji . . . . .	20
4.9	Data TF-IDF . . . . .	20
4.10	Akurasi Tertinggi K-NN . . . . .	21
4.11	Prediksi K-NN pada Data Uji Menggunakan Kelas Bidang Keahlian dengan K= 5 . . . . .	21
4.12	Prediksi K-NN pada Data Uji Menggunakan Kelas Bidang Keahlian dengan K= 10 . . . . .	22
4.13	Prediksi K-NN pada Data Uji Menggunakan Kelas Bidang Keahlian dengan K= 15 . . . . .	22
4.14	Prediksi K-NN pada Data Uji Menggunakan Kelas Bidang Keahlian dengan K= 20 . . . . .	23
4.15	Prediksi K-NN pada Data Uji Menggunakan Kelas Bidang Keahlian dengan K= 25 . . . . .	23
4.16	Prediksi K-NN pada Data Uji Menggunakan 22 Kelas dengan K= 5 . . . . .	24
4.17	Prediksi K-NN pada Data Uji Menggunakan 22 Kelas dengan K= 10 . . . . .	24
4.18	Prediksi K-NN pada Data Uji Menggunakan 22 Kelas dengan K= 15 . . . . .	25
4.19	Prediksi K-NN pada Data Uji Menggunakan 22 Kelas dengan k= 20 . . . . .	25
4.20	Prediksi K-NN pada Data Uji Menggunakan 22 Kelas dengan K= 25 . . . . .	26
4.21	Akurasi Tertinggi <i>K-Nearest Neighbour</i> 3 Kelas . . . . .	26
4.22	Akurasi Tertinggi <i>K-Nearest Neighbour</i> 22 Kelas . . . . .	27



UIN SUSKA RIAU

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim 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.

## DAFTAR SINGKATAN

AWS	: <i>Automatic Weather Station</i>
BSI	: <i>Bisnis Sistem Informasi</i>
IDF	: <i>Inverse Document Frequency</i>
K-NN	: <i>K-Nearest Neighbor</i>
LSA	: <i>Latent Semantic Analysis</i>
MSI	: <i>Manajemen Sistem Informasi</i>
NBC	: <i>Naive Bayes Clasifier</i>
PNN	: <i>Probabilistic Neural Network</i>
RSI	: <i>Rekayasa Sistem Informasi</i>
TA	: <i>Tugas Akhir</i>
TDM	: <i>Text Data Mining</i>
TF	: <i>Term frequency</i>
BRIN	: <i>Badan Riset dan Inovasi Nasional</i>
SINTA	: <i>Science and Technology Index</i>
UIN SUSKA	: <i>Universitas Islam Negeri Sultan Syarif Kasim</i>

### 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.





# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Semakin maju dan mudahnya perkembangan teknologi media penyimpanan digital telah mendorong terjadinya ledakan jumlah dokumen elektronik yang tersimpan dalam *repository* perpustakaan universitas. Berbagai karya ilmiah dari sivitas akademika seperti skripsi, laporan penelitian, dan laporan kerja praktek telah tersedia dalam versi digital. Tetapi, umumnya fenomena ini tidak disertai dengan pertumbuhan jumlah informasi atau pengetahuan yang dapat disarikan dari dokumen-dokumen elektronik tersebut. Agar dapat mengatasi masalah tersebut metode *Text Mining* dibuat dari pengembangan dari metode *data mining*. Algoritma-algoritma dalam *text mining* dibuat untuk dapat mengenali data yang sifatnya semi terstruktur misalnya sinopsis, abstrak maupun isi dari dokumen-dokumen (Akromunnisa dan Hidayat, 2019).

Adapun penelitian ilmiah ialah rangkaian pengamatan yang sambung-mbersambung, berakumulasi, dan melahirkan teori-teori yang mampu menjelaskan dan meramalkan fenomena-fenomena (Rakhmat, 2004). Karya tulis ilmiah merupakan karya tulis yang diakui dalam bidang ilmu pengetahuan, teknologi atau seni yang ditulis atau dikerjakan sesuai dengan tata cara ilmiah dengan mengikuti pedoman dan konvensi ilmiah yang telah disepakati. Di samping itu karya tulis ilmiah merupakan sarana untuk mengutarakan nilai-nilai praktis ataupun nilai-nilai teoretis hasil-hasil pengkajian dan penelitian yang dikerjakan oleh mahasiswa dan dosen (LPM, 2013).

Dalam karya ilmiah terdapat suatu bidang ilmu pengetahuan yang ditulis untuk menyajikan penelitian. *Scientific paper Google Scholar* dosen Sistem Informasi Universitas Islam Negeri Riau (UIN SUSKA Riau) dapat digunakan untuk menentukan klasifikasi judul *paper* dosen sistem informasi berdasarkan bidang ilmu penelitian dosen dengan menggunakan metode K-NN. Pada Program Studi Sistem Informasi UIN SUSKA Riau terdapat dosen yang memiliki profil Google Scholar dan mempublikasikan paper yang tidak sesuai dengan bidangnya. Tidak baik apabila dosen memiliki publikasi yang tidak sesuai dengan bidangnya, berdasarkan undang-undang Kemenristek Badan Riset dan Inovasi Nasional (BRIN) 2019 tentang dan bidang penelitian di *Science and Technology Index* (SINTA).

Pada *Google Scholar* ada beberapa penelitian dosen yang tidak sesuai bidangnya. Oleh karena itu, akan dilakukan *preprocessing* untuk menentukan bidang

Hak Cipta Dilindungi Undang-Undang

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

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



UIN SUSKA RIAU

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

dosen dengan penelitiannya, data yang sudah dilakukan *preprocessing* nantinya dijadikan sebagai data set. Penelitian ini bertujuan untuk mengklasifikasikan *Scientific Paper Google Scholar* dosen Sistem Informasi sesuai dengan bidang ilmunya. Untuk penyetaraan penelitian dosen dengan bidang ilmunya, melihat kecocokan judul paper dengan bidang ilmu dosen tersebut maka dilakukan *text mining* menggunakan algoritma K-NN. Dengan menggunakan K-NN atribut yang digunakan dari *text, unstruktural text* (judul paper) dan kelas target berdasarkan kelompok bidang ilmu dan mata kuliah atau keahlian dosen akan dapat mengklasifikasikan *scientific paper google scholar* dosen Sistem Informasi.

*Google Scholar* adalah bagian dari indeks pencarian *Google* yang lebih besar, yang berisi artikel jurnal teks lengkap, laporan teknis, pracetak, tesis, buku, dan dokumen lainnya, termasuk halaman *Web* terpilih yang dianggap 'ilmiah'. Meskipun *Google Cendekia* mencakup berbagai bidang topikal, tampaknya menjadi yang terkuat dalam ilmu, khususnya kedokteran, dan yang kedua dalam ilmu sosial. Perusahaan ini mengklaim memiliki konten teks lengkap dari semua penerbit utama kecuali *Elsevier* dan *American Chemical Society*, serta layanan *hosting* seperti *Highwire* dan *Ingenta* (Vine, 2006). *Google Cendekia* telah disambut dengan antusiasme dan kritik sejak diperkenalkan pada tahun 2004. Mesin pencarian ini menyediakan cara sederhana untuk mengakses makalah, tesis, buku, abstrak, dan artikel dari situs penerbit akademis, masyarakat profesional, repositori pracetak, universitas dan organisasi ilmiah lainnya (Shultz, 2007).

Definisi *Data mining* sebagai proses untuk memperoleh informasi yang berguna dari *database* yang besar. *Data mining* juga dipahami sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang digunakan untuk membantu pengambilan keputusan. *Data Mining* ialah analisis otomatis dari data yang berjumlah besar atau kompleks yang memiliki tujuan untuk mendapatkan pola atau kecenderungan yang penting yang biasanya tidak disadari keberadaannya. *Data Mining* merupakan analisis dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, *statistic*, *database*, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari *database* yang besar (Bangun, 2019).

Adapun metode yang digunakan adalah Klasifikasi. Klasifikasi merupakan suatu proses kedalam kelas berdasarkan kemiripan sifat yang didalamnya terdapat *data training* sebagai patokan dari klasifikasi serta data uji untuk pengujian. Bertujuan untuk memprediksi objek dari kelas yang belum mempunyai label atau kecenderungan data yang belum terklasifikasi. *Nearest neighbour* adalah suatu pendekatan dalam mencari pola dengan menghitung kedekatan dengan mencocokkan bobot dari

fitur yang ada, pada algoritma ini jarak antara data dapat dihitung yang terdapat nilai K didalamnya. Nilai K pada *Nearest Neighbour* berarti K-data terdekat dari data uji (Efendi dan Mustakim, 2017).

Algoritma K-NN ialah algoritma yang digunakan untuk melakukan klasifikasi terhadap suatu objek, berdasarkan k buah data latih yang jaraknya paling dekat dengan objek tersebut. Syarat nilai K ialah tidak boleh lebih besar dari jumlah data latih, dan nilai K harus ganjil dan lebih dari satu. Dekat atau jauhnya jarak data latih yang paling dekat dengan objek yang akan diklasifikasi dapat dihitung dengan menggunakan metode *cosine similarity* (Rivki dan Bachtiar, 2017). Tujuan dari algoritma ini ialah untuk mengklasifikasikan obyek berdasarkan atribut dan *training sample*. *Classifier* tidak menggunakan apapun untuk dicocokkan dan hanya berdasarkan pada memori. Diberikan titik *query*, akan ditemukan sejumlah K obyek atau (titik *training*) yang paling dekat dengan titik *query*. Klasifikasi menggunakan voting terbanyak diantara klasifikasi dari k obyek. Algoritma K-NN menggunakan klasifikasi ketetangaan sebagai nilai prediksi dari *query instance* yang baru (Lidya, Sitompul, dan Efendi, 2015).

Penelitian-penelitian terkait tentang klasifikasi yang telah dilakukan oleh beberapa peneliti diantaranya (Akromunnisa dan Hidayat, 2019) klasifikasi dokumen Tugas Akhir (skripsi) menggunakan K-NN dengan hasil penelitian menunjukkan klasifikasi menggunakan metode K-NN bisa digunakan untuk mengklasifikasi data intisari bahasa Indonesia dan judul dengan akurasi yang lebih besar tanpa melalui proses *stemming*. (Claudy, Perdana, dan Fauzi, 2018) menggunakan K-NN untuk klasifikasi dokumen *twitter* untuk mengetahui karakter calon karyawan dengan hasil menunjukkan algoritma K-NN dapat diimplementasikan pada sistem klasifikasi kepribadian atau karakter calon karyawan. (Efendi dan Mustakim, 2017) melakukan *Text Mining Classification* Sebagai Rekomendasi Dosen Pembimbing Tugas Akhir Program Studi Sistem Informasi dengan hasil menunjukkan bahwa *Text Mining* dengan algoritma klasifikasi yaitu K-NN untuk rekomendasi dosen pembimbing Tugas Akhir berhasil dilakukan dengan akurasi sebesar 60% dengan menggunakan 98 data latih dan 10 data uji. (Prakasa dan Lhaksmana, 2018) melakukan penelitian terhadap klasifikasi teks dengan menggunakan algoritma K-NN pada kasus kinerja pemerintah di *twitter* dengan hasil dalam proses K-NN atau lebih tepatnya klasifikasi semakin banyak data latih yang digunakan maka akurasi yang didapat semakin bagus karena akan banyak proses pembelajaran pada data latih.

Pada penelitian (Rizaldi dan Mustakim, 2020) tentang Perbandingan Teknik Pembagian Data untuk Klasifikasi Sarana Akses Air pada Algoritma K-NN dan

*Naïve Bayes Classifier* dengan hasil perbandingan teknik pembagian data terbaik menghasilkan *K-Medoid* secara berurutan pada algoritma K-NN pada parameter  $K=10$  dengan nilai akurasi 89.39%, nilai recall sebesar 79.44% dan nilai presisi sebesar 81.77% dan *Naïve Bayes Classifier* dengan nilai akurasi 40.91%, nilai recall sebesar 55.19% dan nilai presisi sebesar 45.69%. Selanjutnya penelitian oleh (Asril, Mustakim, dan Kamila, 2019) mengenai Klasifikasi Dokumen Tugas Akhir Berbasis *Text Mining* menggunakan Metode *Naïve Bayes Classifier* dan K-NN dengan hasil dari percobaan 3 kelas dan 16 kelas diperoleh akurasi terbaik pada percobaan 3 kelas dengan nilai 86.11% untuk NBC dan 91.67% untuk K-NN dan percobaan 16 kelas dengan nilai 25.20% untuk NBC dan 22.05% untuk K-NN. Pada penelitian tersebut menunjukkan hasil dari nilai akurasi K-NN lebih tinggi dibanding dengan NBC.

Adapun penelitian selanjutnya oleh (Hidayat, Mustakim, Fauzi, dan Syukra, 2019) Implementasi Algoritma K-NN dan *Probabilistic Neural Network* untuk Analisis Opini Masyarakat Terhadap Toko Online di Indonesia dengan hasil percobaan menggunakan metode pembagian data dengan 10 K pada *K-Fold Cross Validation* didapat hasil perbandingan akurasi antara K-NN dan PNN pada data Lazada dan Bilibli, dimana pada data Lazada akurasi K-NN lebih tinggi dibandingkan PNN dengan akurasi K-NN sebesar 71.57% sedangkan PNN sebesar 66.71%. Untuk data Bilibli akurasi K-NN juga lebih tinggi dibandingkan PNN dengan akurasi K-NN sebesar 68.29% sedangkan PNN sebesar 65.29%. Jadi hasil akurasi pada data Lazada dan Bilibli menggunakan algoritma K-NN memiliki performa lebih baik dibanding PNN.

Dari latar belakang di atas, akan dilakukan pemodelan Algoritma K-NN dengan judul penelitian “Klasifikasi Dokumen *Scientific Paper Google Scholar* Dosen Sistem Informasi dengan Metode *K-Nearest Neighbor*”.

## 1.2 Perumusan Masalah

Menurut latar belakang yang telah dijabarkan diatas, maka didapat rumusan masalah dari penelitian ini ialah “Bagaimana penerapan algoritma K-NN untuk klasifikasi *scientific paper* dosen Sistem Informasi”.

## 1.3 Batasan Masalah

Batasan masalah Tugas Akhir ini adalah:

1. Data yang digunakan adalah *scientific paper Google Scholar* dosen Sistem Informasi UIN SUSKA Riau.
2. Algoritma klasifikasi yang digunakan yaitu K-NN.
3. Atribut yang akan digunakan untuk klasifikasi adalah profil *Google Scholar*,

judul paper, bidang keahlian, bidang ilmu, dan penulis.

#### 1.4 Tujuan

Tujuan dari penelitian ini ialah untuk mengklasifikasikan dokumen *scientific paper Google Scholar* dosen Sistem Informasi sebagai penyetaraan penelitian dosen dengan bidang ilmunya menggunakan algoritma K-NN.

#### 1.5 Manfaat

Adapun manfaat penelitian ini adalah:

1. Membantu klasifikasi *scientific paper Google Scholar* dosen Sistem Informasi.
2. Membantu penyetaraan penelitian dosen Sistem Informasi sesuai dengan bidang ilmunya.

#### 1.6 Sistematika Penulisan

Sistematika penulisan laporan adalah sebagai berikut:

##### **BAB 1. PENDAHULUAN**

BAB 1 pada Tugas Akhir ini berisi tentang: (1) latar belakang masalah; (2) rumusan masalah; (3) batasan masalah; (4) tujuan; (5) manfaat; dan (6) sistematika penulisan.

##### **BAB 2. LANDASAN TEORI**

BAB 2 pada Tugas Akhir ini berisi tentang: Bab ini berisi uraian mengenai teori-teori umum hingga teori-teori tentang topik yang dibahas, berasal dari buku, jurnal, artikel untuk digunakan dalam landasan teori Tugas Akhir.

##### **BAB 3. METODOLOGI PENELITIAN**

BAB 3 pada Tugas Akhir ini berisi tentang: Bab ini membahas tentang metodologi penelitian dan metodologi dalam menganalisa perhitungan yang digunakan dalam penyusunan Tugas Akhir.

##### **BAB 4. ANALISIS DAN HASIL**

BAB 4 pada Tugas Akhir ini berisi tentang: Bab ini berisi analisa dan hasil tentang perhitungan data menggunakan K-NN.

##### **BAB 5. PENUTUP**

BAB 5 pada Tugas Akhir ini berisi tentang: Bab ini berisi tentang kesimpulan dan saran dari penulisan laporan Tugas Akhir.

## BAB 2

# LANDASAN TEORI

### 2.1 *Text Mining*

*Text mining* atau yang biasa dikenal dengan *text Data Mining* atau pencarian pengetahuan dibasis data textual ialah sebuah proses untuk melakukan pencarian pengetahuan yang berfokus pada data yang berbentuk teks, yang bertujuan untuk mengekstrak informasi yang bermanfaat dan mengidentifikasinya. *Text mining* mempunyai kemiripan dengan *Data Mining*. Keduanya memiliki tujuan yang sama yaitu untuk mendapatkan pengetahuan dan informasi dari sekumpulan data yang besar. Yang menjadi perbedaan antara *Text Mining* dengan *Data Mining* ialah pada *Data Mining* masukan data terstruktur sedangkan dengan *text mining* masukan datanya tidak terstruktur (Rivki dan Bachtiar, 2017).

*Text mining* ialah suatu teknologi yang digunakan untuk menemukan suatu pengetahuan yang bermanfaat dalam suatu koleksi dokumen teks agar didapatkan tren, pola, atau kecocokan teks bahasa alamiah yang dapat digunakan untuk tujuan tertentu. *Text mining* ialah proses penggalian, valid, dan dapat ditindaklanjuti pengetahuan yang tersebar di seluruh dokumen dan memanfaatkan pengetahuan ini untuk lebih mengorganisir informasi sebagai referensi dimasa yang akan datang. *Text mining*, atau juga disebut *Text Data Mining* (TDM), ialah penemuan oleh komputer era baru, informasi yang sebelumnya tidak diketahui, secara otomatis dengan mengekstrasi informasi dari sumber daya yang datanya tidak terstruktur (Wahyudi, 2019). *Text mining* adalah satu langkah dari analisis teks yang dilakukan secara otomatis oleh komputer untuk menggali informasi yang berkualitas terangkum dalam sebuah dokumen dari berbagai rangkaian teks. *Text mining* ialah teknik yang digunakan untuk menangani masalah klasifikasi, *clustering*, *information extraction* dan *information retrieval* (Asril dkk., 2019).

*Data Mining* ialah suatu sebutan yang digunakan untuk menguraikan penemuan pengetahuan didalam basis data. *Data Mining* ialah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan *mechine learning* untuk mengekstraksi dan mengidentifikasi informasi yang berguna dan pengetahuan yang terkait dari berbagai *database* besar (Mustakim dan Oktaviani, 2016).

### 2.2 *Text Preprocessing*

Tahapan *Text Preprocessing* dibutuhkan untuk membersihkan sumber data dari data yang tidak diperlukan (Rivki dan Bachtiar, 2017). *Text Preprocessing* ialah tahapan awal dalam pengolahan data input sebelum memasuki proses taha-

pan utama dari metode *Latent Semantic Analysis (LSA)*. *Preprocessing text* bertujuan untuk penyeragaman dan kemudahan pembacaan serta proses LSA selanjutnya. Dalam *preprocessing* terdapat beberapa tahapan. Adapun tahapan *preprocessing*, yaitu: *case folding*, *tokenizing / parsing*, *filtering*, *stemming*. Berikut penjelasan empat tahapan dalam proses *preprocessing* ialah sebagai berikut:

1. *Case folding* ialah tahapan untuk mengubah semua huruf dalam dokumen menjadi huruf kecil dan karakter selain huruf dihapus dan dianggap *delimiter* (pembatas).
2. *Tokenizing/parsing* ialah tahap pemotongan *string input* berdasarkan tiap kata yang menyusunnya. Bersama dengan, spasi digunakan untuk memisahkan antar kata tersebut.
3. *Filtering* ialah tahapan mengambil kata-kata penting dari hasil *tokenizing*. Proses *filtering* dapat menggunakan algoritma *stoplist* (membuang kata yang kurang penting) atau *wordlist* (menyimpan kata penting). *Stoplist / stopword* ialah kata-kata yang tidak deskriptif yang dapat dihapus dalam pendekatan *bag-of-words*. Contoh *stopword* adalah “yang”, “dan”, “di”, “dari” dan lain-lain.
4. *Stemming* Teknik ini dibutuhkan selain untuk mengurangi jumlah indeks yang berbeda dari suatu dokumen, juga untuk mengelompokan kata-kata lain yang mempunyai kata dasar serta arti yang mirip namun mempunyai bentuk atau *form* yang berbeda karena mendapatkan imbuhan yang berbeda. Sebagai contoh kata bersama, kebersamaan, menyamai, akan distem ke *root word*-nya yaitu “sama” (Prakasa dan Lhaksmana, 2018).

### 2.3 Term Frequency–Inverse Document Frequency (TF-IDF)

TF-IDF ialah singkatan dari *Term Frequency Inverted Document Frequency*. *Term frequency* (TF) ialah frekuensi dari kemunculan suatu kata dalam dokumen yang bersangkutan sementara itu *Inverse Document Frequency* (IDF) ialah suatu pembobotan statistik *global* yang mengkararakteristikan suatu kata dalam keseluruhan koleksi dokumen. IDF merupakan suatu perhitungan dari bagaimana kata didistribusikan secara luas pada koleksi. Nilai IDF berbanding terbalik dengan tingkat kemunculan suatu kata dalam koleksi. Semakin sering suatu kata muncul dalam koleksi, maka nilai IDF dari kata tersebut semakin kecil. Begitu juga sebaliknya. Nilai bobot TF-IDF didefinisikan pada Persamaan 2.1 sebagai berikut:

$$w(t_i, d_j) = \frac{\text{count}(t_i, d_j) \cdot \log(|\text{corpus}|)}{\text{count\_doc}(t_i, \text{corpus})} \quad (2.1)$$

Dimana  $w(t_i)$ , ialah nilai bobot TF-IDF dari kata  $i$  pada dokumen  $j$ ,  $count(t_i, d_j)$  merupakan nilai bobot TF-IDF dari kata  $i$  dalam dokumen  $j$ ,  $corpus$  ialah jumlah seluruh dokumen yang ada dalam koleksi dan  $count\_doc(t_i, corpus)$  ialah jumlah dokumen dalam koleksi yang mengandung kata  $i$ . Penghitungan bobot dari kata tertentu dalam sebuah dokumen dengan menggunakan TF-IDF menunjukkan bahwa deskripsi terbaik dari dokumen adalah kata yang banyak muncul dalam dokumen tersebut dan sangat sedikit muncul pada dokumen yang lain (Sa'dyah dan Ulinuha, 2013).

## 2.4 Klasifikasi

Klasifikasi ialah suatu aktivitas untuk memasukkan sebuah objek ke dalam kelas tertentu dari total jumlah kelas yang tersedia (Rustiyan dan Mustakim, 2017). Klasifikasi ialah suatu proses kedalam kelas berdasarkan kesamaan sifat yang didalamnya terdapat data latih sebagai acuan klasifikasi serta data uji untuk pengujian. Bertujuan untuk memprediksi objek dari suatu kelas yang belum mempunyai label atau kecenderungan data yang belum terklasifikasi. K-NN adalah suatu pendekatan dalam mencari kasus dengan menghitung kedekatan berdasarkan pencocokan bobot dari fitur yang ada, pada algoritma ini jarak antar data dapat dihitung dari data yang terdapat nilai K didalamnya. Nilai K pada K-NN berarti K-data terdekat dari data uji (Efendi dan Mustakim, 2017).

## 2.5 K-Nearest Neighbor

Algoritma K-NN ialah sebuah teknik untuk mengklasifikasi objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Ketepatan algoritma K-NN ini sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan, atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap klasifikasi. Riset terhadap algoritma ini sebagian besar membahas bagaimana memilih dan memberi bobot terhadap fitur, agar performa klasifikasi menjadi lebih baik (Dewi, 2016).

Prinsip kerja dari K-NN ialah mencari jarak dari dua titik yaitu titik *training* dan titik *testing*. Selanjutnya dilakukan evaluasi dengan K tetangga terdekatnya dalam data latih. Jarak yang diaplikasikan ialah persamaan jarak *Euclidean Distance* (Yusra dan Vitriani, 2016). Pada penelitian yang digarap oleh (Efendi dan Mustakim, 2017) K-NN mengklasifikasikan suatu objek berlandaskan data latih yang jaraknya paling dekat dari objek tersebut. Dekat atau jauhnya tetangga umumnya dihitung berdasarkan jarak *euclidean* dengan Persamaan 2.2 sebagai berikut:



$$d = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad (2.2)$$

Keterangan:

d = jarak

a = data uji/testing

b = sampel data

i = variable data

n = Dimensi data

Prinsip kerja dari K-NN ialah mencari jarak antara dua titik yaitu titik *training* dan titik *testing*. Selanjutnya dilakukan evaluasi dengan K tetangga terdekatnya dalam data latih. Jarak yang digunakan ialah persamaan jarak *Euclidean Distance* (Claudy dkk., 2018).

## 2.6 Confusion Matrix

Menurut (Faiza, 2009) *Confusion Matrix* merupakan sebuah model evaluasi klasifikasi berdasarkan data uji dan seluruh data yang diprediksi dengan proporsi yang tepat. Seperti pada Tabel 2.1.

**Tabel 2.1.** Tabel Confusion Matrix 2 Kelas

Classification	Prediction Class	
	Yes	No
Yes	A (True Positive)/TP	B (False Postive)/FN
No	C (False Postive)/FP	D (True Negative)/TN

Dalam penelitian (Adriani, Asian, Nazief, Tahaghoghi, dan Williams, 2007) Terdapat perhitungan tingkat akurasi pada Confusion Matriks 2 kelas berdasarkan Persamaan 2.3 adalah:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} = \frac{A + D}{A + B + C + D} \quad (2.3)$$

## 2.7 Karya ilmiah

Penelitian Ilmiah ialah rangkaian pengamatan yang sambung-bersambung, berakumulasi, dan memunculkan teori-teori yang mampu menjelaskan dan mengamalkan fenomena-fenomena (Rakhmat, 2004). Karya tulis ilmiah adalah karya tulis yang telah diakui dalam bidang ilmu pengetahuan, teknologi atau seni yang ditulis atau dikerjakan berdasarkan tata cara ilmiah dengan mengikuti pedoman dan konvensi ilmiah yang telah disepakati. Di samping itu karya tulis ilmiah merupakan wahana untuk menyajikan nilai-nilai praktis dan juga nilai-nilai teoritis hasil-



pengkajian dan penelitian yang dikerjakan oleh mahasiswa dan dosen (LPM, 2013)

## 2.8 Google Scholar atau Cendekia

*Google Scholar* atau *Google Cendekia* adalah bagian dari indeks pencarian *Google* yang lebih besar, yang terdiri dari artikel jurnal teks lengkap, laporan teknis, pracetak, tesis, buku, dan dokumen lainnya, termasuk halaman *Web* terpilih yang dianggap 'ilmiah'. Meskipun *Google Cendekia* mencakup berbagai bidang topikal, tampaknya menjadi yang terkuat dalam ilmu, khususnya kedokteran, dan yang kedua dalam ilmu sosial. Perusahaan ini mengklaim memiliki konten teks lengkap dari semua penerbit utama kecuali *Elsevier* dan *American Chemical Society*, serta layanan *hosting* seperti *Highwire* dan *Ingenta* (Vine, 2006). *Google Cendekia* telah disambut dengan antusiasme dan kritik sejak diperkenalkan pada tahun 2004. Mesin pencarian ini menyediakan cara sederhana untuk mengakses makalah, tesis, buku, abstrak, dan artikel dari situs penerbit akademis, masyarakat profesional, repositori pracetak, universitas dan organisasi ilmiah lainnya (Shultz, 2007).

## 2.9 Penelitian Terdahulu

Penelitian yang dikerjakan oleh Akromunisa dan Hidayat (2019) adalah Penerapan *text mining* dalam klasifikasi dokumen Tugas Akhir (TA) menggunakan K-NN. Penelitian tersebut berguna untuk melakukan klasifikasi dokumen TA secara otomatis. Hasil yang diperoleh yaitu klasifikasi menggunakan metode K-NN dapat digunakan untuk mengklasifikasi data intisari bahasa Indonesia dan judul dengan akurasi yang lebih tinggi tanpa melalui proses *stemming* (Akromunnisa dan Hidayat, 2019).

Adapun penelitian yang dilakukan oleh Efendi dan Mustakim ditahun 2017 tentang klasifikasi *text mining* sebagai rekomendasi dosen pembimbing TA Program Studi Sistem Informasi dengan algoritma K-NN. Nilai k yang digunakan yaitu 15 dengan akurasi yang diperoleh sebesar 60% (Efendi dan Mustakim, 2017). Selanjutnya penelitian yang dilakukan oleh Claudy dkk pada tahun 2018 tentang klasifikasi dokumen *Twitter* untuk mengetahui karakter calon karyawan. Kesimpulan dari penelitian tersebut yaitu hasil menunjukkan algoritma K-NN dapat diaplikasikan pada sistem klasifikasi kepribadian atau karakter calon karyawan (Claudy dkk., 2018).

Selanjutnya penelitian yang dilakukan oleh Prakasa dan Lhaksamana pada tahun 2018 tentang klasifikasi teks dengan algoritma K-NN pada kasus kinerja pemerintah di *Twitter*. Kesimpulan dari penelitian ini adalah dalam proses K-NN semakin banyak data latih yang digunakan maka akurasi yang diperoleh semakin baik karena akan banyak proses pembelajaran pada data *training* (Prakasa dan Lhaksamana, 2018). Terdapat pula penelitian oleh Puspita tahun 2021 tentang komparasi

metode K-NN, *Decision Tree*, dan *Naïve Bayes* Terhadap Analisis Sentimen Pengguna Layanan BPJS, tujuan dari penelitian yang dilakukan ini ialah untuk mengetahui tingkat, akurasi, dari, tiga metode, yang berbeda diantaranya K-NN, *Decision Tree* dan *Naïve Bayes*, dengan menggunakan *tools RapidMiner*, pada penelitian ini, diketahui bahwa metode *Decision Tree* merupakan metode yang tingkat akurasinya lebih besar dibandingkan kedua metode lainnya dengan tingkat akurasi sebesar 96.13% (Puspita dan Widodo, 2021). Adapun penelitian yang dilakukan oleh Noviansyah tahun 2018 tentang Penerapan *Data Mining* Menggunakan Metode K-NN Untuk Klasifikasi Indeks Cuaca Kebakaran Berdasarkan Data *Automatic Weather Station (AWS)* dengan hasil pengujian menggunakan 252 data uji, memperoleh akurasi persentase keberhasilan 80.16% dengan nilai K=5 (Noviansyah, Rismawan, dan Midyanti, 2018).

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

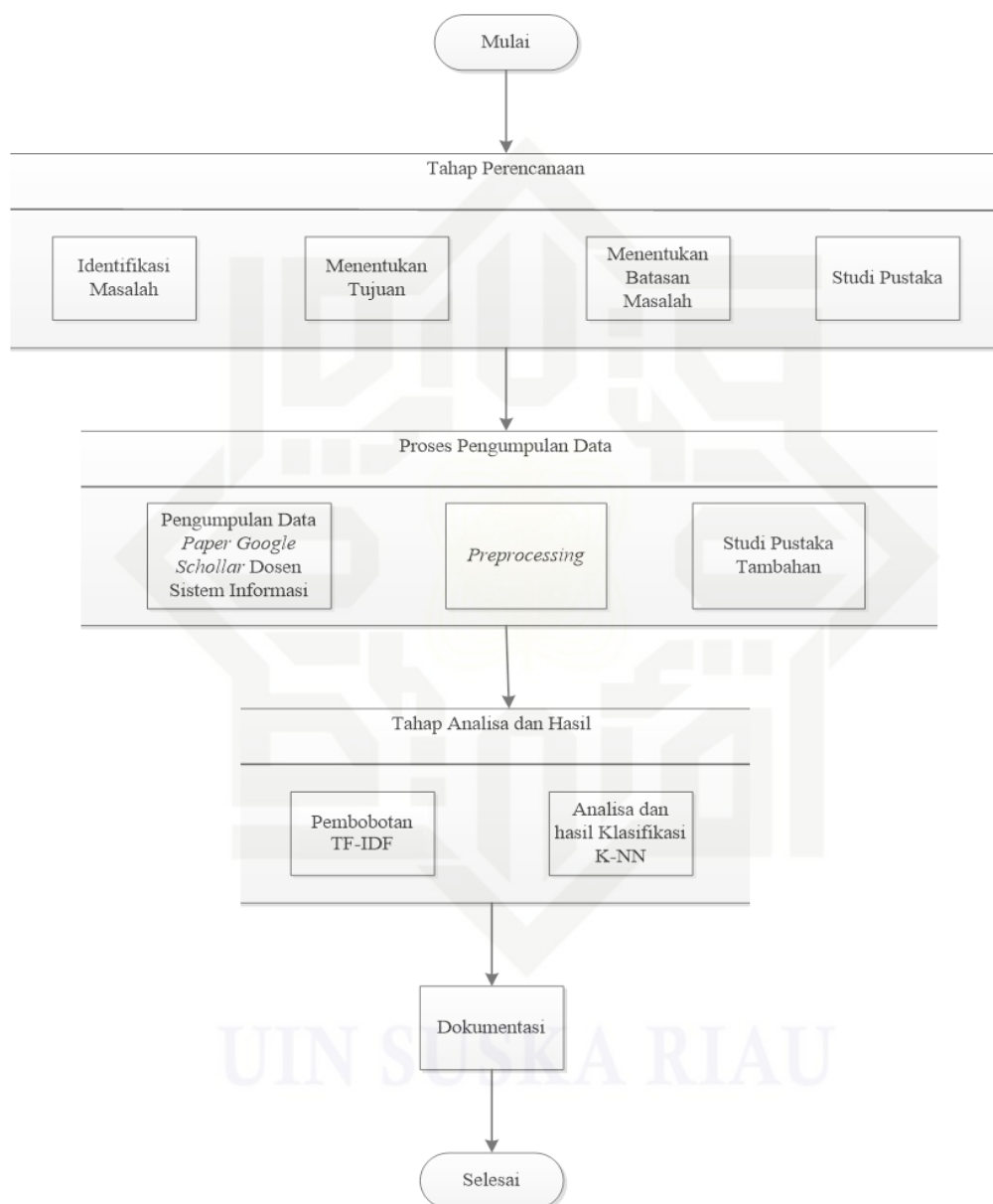
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan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



## BAB 3 METODOLOGI PENELITIAN

### 3.1 Proses Alur Penelitian

Terdapat beberapa langkah dalam penyusunan Tugas Akhir ini dapat dilihat pada Gambar 3.1.



**Gambar 3.1.** Metodologi Penelitian

### 3.2 Tahapan Perencanaan

Tahap perencanaan ialah tahapan yang harus dilakukan saat akan memulai penelitian, data yang direncanakan adalah:

Hak Cipta Dilindungi Undang-Undang

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

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.

1. Identifikasi Masalah  
Mengamati dan mengidentifikasi permasalahan yang terdapat pada *Scientific Paper Google Scholar* dosen Sistem Informasi UIN SUSKA Riau.
2. Menentukan Tujuan Penelitian  
Tujuan dari penelitian ini ialah untuk penyetaraan penelitian dosen sesuai dengan bidang ilmunya di Jurusan Sistem Informasi UIN SUSKA Riau.
3. Menentukan Batasan Masalah  
Menentukan batasan masalah berguna agar dapat berfokus pada objek penelitian sehingga tidak melenceng dari lingkup objek penelitian yang telah ditetapkan.
4. Studi Pustaka  
Studi pustaka dilaksanakan untuk mencari rujukan-rujukan utama yang berkaitan terhadap penelitian ini.

### 3.3 Proses Pengumpulan Data

Pada tahap ini akan dilaksanakan pengumpulan data yang akan dipakai dalam penelitian ini. Tahap dari proses pengumpulan data ialah seperti berikut:

1. Pengumpulan Data *Paper Google scholar* Dosen Sistem Informasi  
Pengumpulan data dilakukan dengan mencari judul *paper* dosen sistem informasi pada profil *Google Scholar*. Data terakhir diambil pada desember 2020 dengan jumlah data sebanyak 362 data.
2. *Preprocessing*  
Berikut ini ialah tahapan dari *preprocessing*:
  - (a) *Tokenizing*  
Tahapan *tokenizing*, ialah tahap pemisahan kata berdasarkan karakter spasi pada setiap dokumen.
  - (b) *Filtering*  
Tahapan *filtering*, adalah tahap menghapus kata-kata yang tidak diperlukan dari proses *tokenizing*. Proses penghapusan kata dilakukan dengan *stopword* atau kamus kata.
  - (c) *Stemming*  
Proses *stemming* ialah proses yang mengubah bentuk dari suatu kata menjadi bentuk kata dasarnya.
3. Studi Pustaka Tambahan  
Studi pustaka lanjutan dilakukan untuk mencari referensi tambahan yang diperlukan oleh peneliti. Sumber yang digunakan dalam menyelesaikan penelitian ini berasal dari jurnal, buku, skripsi dan *website* resmi.

### 3.4 Tahap Analisa dan Hasil

1. Pembobotan TF-IDF

Pembobotan yang digunakan pada penelitian ini yaitu TF-IDF, pembobotan dilakukan karena setiap kata memiliki tingkat kepentingan yang berbeda dalam setiap dokumen.

2. Pada tahap ini dilakukan klasifikasi menggunakan algoritma (K-NN), kemudian dilanjutkan dengan menghitung akurasi algoritma.

### 3.5 Tahap Dokumentasi

Tahap dokumentasi adalah tahap mendokumentasikan hasil dari penelitian yang telah dilakukan dengan pembuatan laporan. Hasil akhir pada tahapan ini berwujud dokumen laporan Tugas Akhir.

#### 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.

## BAB 5

### PENUTUP

#### 5.1 Kesimpulan

Akurasi yang diperoleh pada perhitungan K-NN yang diterapkan dalam klasifikasi dokumen *scientific paper Google Scholar* dosen Sistem Informasi dengan 3 kelas bidang keahlian menghasilkan nilai akurasi terbaik yaitu 70.64% terdapat pada nilai K=15. Selanjutnya pada data uji dengan menggunakan 22 kelas bidang ilmu menghasilkan nilai akurasi terbaik yang sama dengan 3 kelas bidang keahlian dengan nilai K=10. Berdasarkan percobaan tersebut, kesimpulan yang didapat dari hasil pengklasifikasian dokumen *scientific paper Google Scholar* dosen Sistem Informasi yaitu 70.64% paper sesuai dengan bidang keahlian serta bidang ilmu dosen. Jumlah data yang digunakan adalah 253 *data training* dan 109 *data testing*.

#### 5.2 Saran

Pada penelitian ini tentunya terdapat beberapa kekurangan dan kelemahan. Berikut saran yang dapat diberikan:

1. Melakukan percobaan menggunakan algoritma lainnya untuk mendapatkan tingkat akurasi terbaik dari algoritma.
2. Menambah jumlah data uji lebih banyak agar memperoleh nilai akurasi yang lebih tinggi.



## DAFTAR PUSTAKA

- Adriani, M., Asian, J., Nazief, B., Tahaghoghi, S. M., dan Williams, H. E. (2007). Stemming indonesian: A confix-stripping approach. *ACM Transactions on Asian Language Information Processing (TALIP)*, 6(4), 1–33.
- Akromunnisa, K., dan Hidayat, R. (2019). Klasifikasi dokumen tugas akhir (skripsi) menggunakan k-nearest neighbor. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 4(1), 69–75.
- Asril, H., Mustakim, M., dan Kamila, I. (2019). Klasifikasi dokumen tugas akhir berbasis text mining menggunakan metode naïve bayes classifier dan k-nearest neighbor. Dalam *Seminar nasional teknologi informasi komunikasi dan industri* (hal. 178–186).
- Bangun, I. R. B. (2019). Prediksi putusan hukuman tindakan kriminalitas dengan menggunakan algoritma nearest neighbor (studi kasus: Pengadilan negeri lubuk pakam). *Informasi dan Teknologi Ilmiah (INTI)*, 14(1), 77–81.
- Claudy, Y. I., Perdana, R. S., dan Fauzi, M. A. (2018). Klasifikasi dokumen twitter untuk mengetahui karakter calon karyawan menggunakan algoritme k-nearest neighbor (knn). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN, 2548, 964X*.
- Dewi, S. (2016). Komparasi 5 metode algoritma klasifikasi data mining pada prediksi keberhasilan pemasaran produk layanan perbankan. *Techno Nusa Mandiri: Journal of Computing and Information Technology*, 13(1), 60–65.
- Efendi, Z., dan Mustakim, M. (2017). Text mining classification sebagai rekomendasi dosen pembimbing tugas akhir program studi sistem informasi. Dalam *Seminar nasional teknologi informasi komunikasi dan industri* (hal. 235–242).
- Faiza, N. N. (2009). Prediksi tingkat keberhasilan mahasiswa tingkat i ipb dengan metode k-nearest neighbor.
- Hidayat, A., Mustakim, M., Fauzi, M. Z., dan Syukra, I. (2019). Implementasi algoritma k-nearest neighbor dan probabilistic neural network untuk analisis opini masyarakat terhadap toko online di indonesia. Dalam *Seminar nasional teknologi informasi komunikasi dan industri* (hal. 254–261).
- Lidya, S. K., Sitompul, O. S., dan Efendi, S. (2015). Sentiment analysis pada teks bahasa indonesia menggunakan support vector machine (svm) dan k-nearest neighbor (k-nn). Dalam *Seminar nasional teknologi informasi dan komunikasi*.
- LPM, T. P. (2013). *Pedoman mutu karya ilmiah mahasiswa*. Suska Press.







- Mustakim, M., dan Oktaviani, G. (2016). Algoritma k-nearest neighbor classification sebagai sistem prediksi predikat prestasi mahasiswa. *SITEKIN: Jurnal Sains, Teknologi dan Industri*, 13(2), 195–202.
- Noviansyah, M. R., Rismawan, T., dan Midyanti, D. M. (2018). Penerapan data mining menggunakan metode k-nearest neighbor untuk klasifikasi indeks cuaca kebakaran berdasarkan data aws (automatic weather station)(studi kasus: kabupaten kubu raya). *Coding Jurnal Komputer dan Aplikasi*, 6(2).
- Prakasa, O. S. Y., dan Lhaksana, K. M. (2018). Klasifikasi teks dengan menggunakan algoritma k-nearest neighbor pada kasus kinerja pemerintah di twitter. *eProceedings of Engineering*, 5(3).
- Puspita, R., dan Widodo, A. (2021). Perbandingan metode knn, decision tree, dan naïve bayes terhadap analisis sentimen pengguna layanan bpjs. *Jurnal Informatika Universitas Pamulang*, 5(4), 646–654.
- Rakhmat, J. (2004). *Metode penelitian komunikasi*. PT REMAJA ROSDAKARYA.
- Rivki, M., dan Bachtiar, A. M. (2017). Implementasi algoritma k-nearest neighbor dalam pengklasifikasian follower twitter yang menggunakan bahasa indonesia. *Jurnal Sistem Informasi*, 13(1), 31–37.
- Rizaldi, S. T., dan Mustakim, M. (2020). Perbandingan teknik pembagian data untuk klasifikasi sarana akses air pada algoritma k-nearest neighbor dan naïve bayes classifier. Dalam *Seminar nasional teknologi informasi komunikasi dan industri* (hal. 130).
- Rustiyan, R., dan Mustakim, M. (2017). Klasifikasi jenis perairan pada kapal perikanan di indonesia menggunakan k-nearest neighbor. Dalam *Seminar nasional teknologi informasi komunikasi dan industri* (hal. 178–185).
- Sa'dyah, H., dan Ulinnuha, N. (2013). Penggunaan fitur abstraksi dan catatan publikasi penulis untuk klasifikasi artikel ilmiah dengan metadata yang terbatas. *Jurnal Ilmiah Teknologi Informasi*, 11(1), 29–33.
- Shultz, M. (2007). Comparing test searches in pubmed and google scholar. *Journal of the Medical Library Association: JMLA*, 95(4), 442.
- Tala, F. (2003). A study of stemming effects on information retrieval in bahasa indonesia.
- Vine, R. (2006). Google scholar. *Journal of the Medical Library Association*, 94(1), 97.
- Wahyudi, M. D. R. (2019). Penerapan algoritma cosine similarity pada text mining terjemah al-qur'an berdasarkan keterkaitan topik. *Semesta Teknika*, 22(1), 41–50.
- Yusra, D. O., dan Vitriani, Y. (2016). Perbandingan klasifikasi tugas akhir maha-

siswa jurusan teknik informatika menggunakan metode naïve bayes classifier dan k-nearest neighbor. *J. Sains, Teknol. dan Ind*, 14(1), 79–85.



UIN SUSKA RIAU

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim 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.



UIN SUSKA RIAU

## LAMPIRAN A

### DAFTAR KATA *STOPWORD*

No	<i>Stopword</i>	No	<i>Stopword</i>
1	ada	380	melainkan
2	adalah	381	melakukan
3	adanya	382	melalui
4	adapun	383	melihat
5	agak	384	melihatnya
6	agakny	385	memang
7	agar	386	memastikan
8	akan	387	memberi
9	akankah	388	memberikan
10	akhir	389	membuat
11	akhiri	390	memerlukan
12	akhirnya	391	memihak
13	aku	392	meminta
14	akulah	393	memintakan
15	amat	394	memisalkan
16	amatlah	395	memperbuat
17	anda	396	mempergunakan
18	andalah	397	memperkirakan
19	antar	398	memperlihatkan
20	antara	399	mempersiapkan
21	antaranya	400	mempersoalkan
22	apa	401	mempertanyakan
23	apaan	402	mempunyai
24	apabila	403	memulai
25	apakah	404	memungkinkan
26	apalagi	405	menaiki
27	apatah	406	menambahkan
28	artinya	407	menandaskan
29	asal	408	menanti
30	asalkan	409	menanti-nanti
31	atas	410	menantikan
32	atau	411	menanya
33	ataukah	412	menanyai
34	ataupun	413	menanyakan
35	awal	414	mendapat
36	awalnya	415	mendapatkan
37	bagai	416	mendatang
38	bagaikan	417	mendatangi
39	bagaimana	418	mendatangkan
40	bagaimanakah	419	menegaskan
41	bagaimanapun	420	mengakhiri
42	bagi	421	mengapa
43	bagian	422	mengatakan
44	bahkan	423	mengatakannya
45	bahwa	424	mengenai
46	bahwasanya	425	mengerjakan
47	baik	426	mengetahui

#### 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 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.

**Tabel A.1** Daftar Kata *Stopwords* (Tabel lanjutan...)

No	Stopword	No	Stopword
48	bakal	427	menggunakan
49	bakalan	428	menghendaki
50	balik	429	mengibaratkan
51	banyak	430	mengibaratkannya
52	bapak	431	mengingat
53	baru	432	mengingatkan
54	bawah	433	menginginkan
55	beberapa	434	mengira
56	begini	435	mengucapkan
57	beginian	436	mengucapkannya
58	beginikah	437	mengungkapkan
59	beginilah	438	menjadi
60	begitu	439	menjawab
61	begitukah	440	menjelaskan
62	begitulah	441	menuju
63	begitupun	442	menunjuk
64	bekerja	443	menunjuki
65	belakang	444	menunjukkan
66	belakangan	445	menunjuknya
67	belum	446	menurut
68	belumah	447	menuturkan
69	benar	448	menyampaikan
70	benarkah	449	menyangkut
71	benarlah	450	menyatakan
72	berada	451	menyebutkan
73	berakhir	452	menyeluruh
74	berakhirilah	453	menyiapkan
75	berakhirnya	454	merasa
76	berapa	455	mereka
77	berapakah	456	merekalah
78	berapalah	457	merupakan
79	berapapun	458	meski
80	berarti	459	meskipun
81	berawal	460	meyakini
82	berbagai	461	meyakinkan
83	berdatangan	462	minta
84	beri	463	mirip
85	berikan	464	misal
86	berikut	465	misalkan
87	berikutnya	466	misalnya
88	berjumlah	467	mula
89	berkali-kali	468	mulai
90	berkata	469	mulailah
91	berkehendak	470	mulanya
92	berkeinginan	471	mungkin
93	berkenaan	472	mungkinkah
94	berlainan	473	nah
95	berlalu	474	naik
96	berlangsung	475	namun
97	berlebihan	476	nanti
98	bermacam	477	nantinya

**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.

**Tabel A.1** Daftar Kata *Stopwords* (Tabel lanjutan...)

No	Stopword	No	Stopword
99	bermacam-macam	478	nyaris
100	bermaksud	479	nyatanya
101	bermula	480	oleh
102	bersama	481	olehnya
103	bersama-sama	482	pada
104	bersiap	483	padahal
105	bersiap-siap	484	padanya
106	bertanya	485	pak
107	bertanya-tanya	486	paling
108	berturut	487	panjang
109	berturut-turut	488	pantas
110	bertutur	489	para
111	berujar	490	pasti
112	berupa	491	pastilah
113	besar	492	penting
114	betul	493	pentingnya
115	betulkah	494	per
116	biasa	495	percuma
117	biasanya	496	perlu
118	bila	497	perlukah
119	bilakah	498	perlunya
120	bisa	499	pernah
121	bisakah	500	persoalan
122	boleh	501	pertama
123	bolehkah	502	pertama-tama
124	bolehlah	503	pertanyaan
125	buat	504	pertanyakan
126	bukan	505	pihak
127	bukankah	506	pihaknya
128	bukanlah	507	pukul
129	bukannya	508	pula
130	bulan	509	pun
131	bung	510	punya
132	cara	511	rasa
133	caranya	512	rasanya
134	cukup	513	rata
135	cukupkah	514	rupanya
136	cukuplah	515	saat
137	cuma	516	saatnya
138	dahulu	517	saja
139	dalam	518	sajalah
140	dan	519	saling
141	dapat	520	sama
142	dari	521	sama-sama
143	daripada	522	sambil
144	datang	523	sampai
145	dekat	524	sampai-sampai
146	demi	525	sampaikan
147	demikian	526	sana
148	demikianlah	527	sangat
149	dengan	528	sangatlah

**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.

**Tabel A.1** Daftar Kata *Stopwords* (Tabel lanjutan...)

No	Stopword	No	Stopword
150	depan	529	satu
151	di	530	saya
152	dia	531	sayalah
153	diakhiri	532	se
154	diakhirinya	533	sebab
155	dialah	534	sebabnya
156	diantara	535	sebagai
157	diantaranya	536	sebagaimana
158	diberi	537	sebagainya
159	diberikan	538	sebagian
160	diberikannya	539	sebaik
161	dibuat	540	sebaik-baiknya
162	dibuatnya	541	sebaiknya
163	didapat	542	sebaliknya
164	didatangkan	543	sebanyak
165	digunakan	544	sebegini
166	diibaratkan	545	sebegini
167	diibaratkannya	546	sebelum
168	diingat	547	sebelumnya
169	diingatkan	548	sebenarnya
170	diinginkan	549	seberapa
171	dijawab	550	sebesar
172	dijelaskan	551	sebetulnya
173	dijelaskannya	552	sebisanya
174	dikarenakan	553	sebuah
175	dikatakan	554	sebut
176	dikatakannya	555	sebutlah
177	dikerjakan	556	sebutnya
178	diketahui	557	secara
179	diketahuinya	558	secukupnya
180	dikira	559	sedang
181	dilakukan	560	sedangkan
182	dilalui	561	sedemikian
183	dilihat	562	sedikit
184	dimaksud	563	sedikitnya
185	dimaksudkan	564	seenaknya
186	dimaksudkannya	565	segala
187	dimaksudnya	566	segalanya
188	diminta	567	segera
189	dimintai	568	seharusnya
190	dimisalkan	569	sehingga
191	dimulai	570	seingat
192	dimulailah	571	sejak
193	dimulainya	572	sejauh
194	dimungkinkan	573	sejenak
195	dini	574	sejumlah
196	dipastikan	575	sekadar
197	diperbuat	576	sekadarnya
198	diperbuatnya	577	sekali
199	dipergunakan	578	sekali-kali
200	diperkirakan	579	sekalian

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.

Tabel A.1 Daftar Kata *Stopwords* (Tabel lanjutan...)

No	Stopword	No	Stopword
201	diperlihatkan	580	sekaligus
202	diperlukan	581	sekalipun
203	diperlukannya	582	sekarang
204	dipersoalkan	583	sekarang
205	dipertanyakan	584	sekecil
206	dipunyai	585	seketika
207	diri	586	sekiranya
208	dirinya	587	sekitar
209	disampaikan	588	sekitarnya
210	disebut	589	sekurang-kurangnya
211	disebutkan	590	sekurangnya
212	disebutkannya	591	sela
213	disini	592	selain
214	disinilah	593	selaku
215	ditambahkan	594	selalu
216	ditandaskan	595	selama
217	ditanya	596	selama-lamanya
218	ditanyai	597	selamanya
219	ditanyakan	598	selanjutnya
220	ditegaskan	599	seluruh
221	ditujukan	600	seluruhnya
222	ditunjuk	601	semacam
223	ditunjuki	602	semakin
224	ditunjukkan	603	semampu
225	ditunjukkannya	604	semampunya
226	ditunjuknya	605	semasa
227	dituturkan	606	semasih
228	dituturkannya	607	semata
229	diucapkan	608	semata-mata
230	diucapkannya	609	semaunya
231	diungkapkan	610	sementara
232	dong	611	semisal
233	dua	612	semisalnya
234	dulu	613	sempat
235	empat	614	semua
236	enggak	615	semuanya
237	enggaknya	616	semula
238	entah	617	sendiri
239	entahlah	618	sendirian
240	guna	619	sendirinya
241	gunakan	620	seolah
242	hal	621	seolah-olah
243	hampir	622	seorang
244	hanya	623	sepanjang
245	hanyalah	624	sepantasnya
246	hari	625	sepantasnyalah
247	harus	626	seperlunya
248	haruslah	627	seperti
249	harusnya	628	sepertinya
250	hendak	629	sepihak
251	hendaklah	630	sering

**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.

**Tabel A.1** Daftar Kata *Stopwords* (Tabel lanjutan...)

No	Stopword	No	Stopword
252	hendaknya	631	seringnya
253	hingga	632	serta
254	ia	633	serupa
255	ialah	634	sesaat
256	ibarat	635	sesama
257	ibaratkan	636	sesampai
258	ibaratnya	637	sesegera
259	ibu	638	sesekali
260	ikut	639	seseorang
261	ingat	640	sesuatu
262	ingat-ingat	641	sesuatunya
263	ingin	642	sesudah
264	inginkah	643	sesudahnya
265	inginkan	644	setelah
266	ini	645	setempat
267	inikah	646	setengah
268	inilah	647	seterusnya
269	itu	648	setiap
270	itukah	649	setiba
271	itulah	650	setibanya
272	jadi	651	setidak-tidaknya
273	jadilah	652	setidaknya
274	jadinya	653	setinggi
275	jangan	654	seusai
276	jangankan	655	sewaktu
277	janganlah	656	siap
278	jauh	657	siapa
279	jawab	658	siapakah
280	jawaban	659	siapapun
281	jawabnya	660	sini
282	jaslah	661	sinilah
283	jelaskan	662	soal
284	jelastah	663	soalnya
285	jelastah	664	suatu
286	jika	665	sudah
287	jikalau	666	sudahkah
288	juga	667	sudahlah
289	jumlah	668	supaya
290	jumlahnya	669	tadi
291	justru	670	tadinya
292	kala	671	tahu
293	kalau	672	tahun
294	kalaulah	673	tak
295	kalaupun	674	tambah
296	kalian	675	tambahnya
297	kami	676	tampak
298	kamilah	677	tampaknya
299	kamu	678	tandas
300	kamulah	679	tandasnya
301	kan	680	tanpa
302	kapan	681	tanya



**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.

**Tabel A.1** Daftar Kata *Stopwords* (Tabel lanjutan...)

No	Stopword	No	Stopword
303	kapankah	682	tanyakan
304	kapapun	683	tanyanya
305	karena	684	tapi
306	karenanya	685	tegas
307	kasus	686	tegasnya
308	kata	687	telah
309	katakan	688	tempat
310	katakanlah	689	tengah
311	katanya	690	tentang
312	ke	691	tentu
313	keadaan	692	tentulah
314	kebetulan	693	tentunya
315	kecil	694	tepat
316	kedua	695	terakhir
317	keduanya	696	terasa
318	keinginan	697	terbanyak
319	kelamaan	698	terdahulu
320	kelihatan	699	terdapat
321	kelihatannya	700	terdiri
322	kelima	701	terhadap
323	keluar	702	terhadapnya
324	kembali	703	teringat
325	kemudian	704	teringat-ingat
326	kemungkinan	705	terjadi
327	kemungkinannya	706	terjadilah
328	kenapa	707	terjadinya
329	kepada	708	terkira
330	kepadanya	709	terlalu
331	kesampaian	710	terlebih
332	keseluruhan	711	terlihat
333	keseluruhannya	712	termasuk
334	keterlalu	713	ternyata
335	ketika	714	tersampaikan
336	khususnya	715	tersebut
337	kini	716	tersebutlah
338	kinilah	717	tertentu
339	kira	718	tertuju
340	kira-kira	719	terus
341	kiranya	720	terutama
342	kita	721	tetap
343	kitalah	722	tetapi
344	kok	723	tiap
345	kurang	724	tiba
346	lagi	725	tiba-tiba
347	lagian	726	tidak
348	lah	727	tidakkah
349	lain	728	tidaklah
350	lainnya	729	tiga
351	lalu	730	tinggi
352	lama	731	toh
353	lamanya	732	tunjuk

**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.

**Tabel A.1** Daftar Kata *Stopwords* (Tabel lanjutan...)

No	Stopword	No	Stopword
354	lanjut	733	turut
355	lanjutnya	734	tutor
356	lebih	735	tuturnya
357	lewat	736	ucap
358	lima	737	ucapnya
359	luar	738	ujar
360	macam	739	ujarnya
361	maka	740	umum
362	makanya	741	umumnya
363	makin	742	ungkap
364	malah	743	ungkapnya
365	malahan	744	untuk
366	mampu	745	usah
367	mampukah	746	usai
368	mana	747	waduh
369	manakala	748	wah
370	manalagi	749	wahai
371	masa	750	waktu
372	masalah	751	waktunya
373	masalahnya	752	walau
374	masih	753	walaupun
375	masihkah	754	wong
376	masing	755	yaitu
377	masing-masing	756	yakin
378	mau	757	yakni
379	maupun	758	yang

sumber : (Tala, 2003)

## LAMPIRAN B

### HASIL PREDIKSI PERCOBAAN KNN 3 KELAS

#### HASIL PREDIKSI PERCOBAAN 3 KELAS K=5

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
1	0.399	0.000	0.601	BSI	BSI
2	0.400	0.402	0.198	RSI	BSI
3	0.198	0.000	0.802	BSI	BSI
4	0.232	0.192	0.576	BSI	BSI
5	0.000	0.000	1.000	BSI	RSI
6	0.000	1.000	0.000	RSI	RSI
7	0.000	0.000	1.000	BSI	MSI
8	0.169	0.168	0.663	BSI	RSI
9	0.000	0.329	0.671	BSI	RSI
10	0.202	0.000	0.798	BSI	MSI
11	0.208	0.397	0.394	RSI	BSI
12	0.590	0.202	0.208	MSI	MSI
13	0.000	0.386	0.614	BSI	RSI
14	0.000	0.410	0.590	BSI	RSI
15	0.413	0.196	0.390	MSI	BSI
16	0.182	0.630	0.188	RSI	MSI
17	0.163	0.000	0.837	BSI	RSI
18	0.000	0.194	0.806	BSI	RSI
19	0.000	1.000	0.000	RSI	RSI
20	0.000	0.177	0.823	BSI	RSI
21	0.000	0.407	0.593	BSI	RSI
22	0.000	0.000	1.000	BSI	RSI
23	0.000	0.000	1.000	BSI	RSI
24	0.000	0.000	1.000	BSI	RSI
25	0.191	0.187	0.622	BSI	BSI
26	0.000	0.220	0.780	BSI	RSI
27	0.000	0.187	0.813	BSI	RSI
28	0.000	0.418	0.582	BSI	RSI
29	0.355	0.000	0.645	BSI	RSI
30	0.000	0.206	0.794	BSI	RSI
31	0.000	0.000	1.000	BSI	RSI
32	0.000	0.240	0.760	BSI	RSI
33	0.198	0.198	0.604	BSI	RSI
34	0.000	0.000	1.000	BSI	RSI
35	0.213	0.187	0.600	BSI	RSI
36	0.000	0.608	0.392	RSI	RSI
37	0.000	0.211	0.789	BSI	RSI
38	0.209	0.393	0.399	BSI	RSI
39	0.600	0.182	0.218	MSI	BSI
40	0.000	1.000	0.000	RSI	MSI
41	0.000	0.410	0.590	BSI	RSI
42	0.000	0.485	0.515	BSI	MSI
43	0.195	0.203	0.602	BSI	RSI
44	0.207	0.193	0.600	BSI	RSI
45	0.000	0.793	0.207	RSI	RSI

#### 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.



Tabel B.1 (Tabel lanjutan...)

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
46	0.000	0.604	0.396	RSI	RSI
47	0.000	0.200	0.800	BSI	RSI
48	0.392	0.204	0.404	BSI	RSI
49	0.619	0.196	0.185	MSI	RSI
50	0.000	0.192	0.808	BSI	RSI
51	0.000	0.000	1.000	BSI	MSI
52	0.414	0.391	0.195	MSI	BSI
53	0.000	0.203	0.797	BSI	RSI
54	0.191	0.187	0.622	BSI	RSI
55	0.198	0.208	0.594	BSI	BSI
56	0.198	0.201	0.601	BSI	BSI
57	0.197	0.220	0.583	BSI	MSI
58	0.409	0.389	0.202	MSI	BSI
59	0.391	0.417	0.192	RSI	MSI
60	0.174	0.826	0.000	RSI	MSI
61	0.000	0.352	0.648	BSI	RSI
62	0.000	0.396	0.604	BSI	MSI
63	0.000	0.000	1.000	BSI	BSI
64	0.000	0.193	0.807	BSI	RSI
65	0.175	0.000	0.825	BSI	RSI
66	0.000	0.808	0.192	RSI	MSI
67	0.000	0.173	0.827	BSI	RSI
68	0.000	0.000	1.000	BSI	BSI
69	0.000	0.605	0.395	RSI	RSI
70	0.000	0.186	0.814	BSI	RSI
71	0.000	0.377	0.623	BSI	RSI
72	0.000	0.395	0.605	BSI	RSI
73	0.424	0.000	0.576	BSI	BSI
74	0.223	0.000	0.777	BSI	BSI
75	0.000	0.449	0.551	BSI	MSI
76	0.000	0.189	0.811	BSI	RSI
77	0.188	0.223	0.589	BSI	RSI
78	0.000	0.200	0.800	BSI	RSI
79	0.000	0.380	0.620	BSI	RSI
80	0.184	0.000	0.816	BSI	RSI
81	0.196	0.209	0.595	BSI	RSI
82	0.209	0.000	0.791	BSI	MSI
83	0.000	0.809	0.191	RSI	MSI
84	0.000	0.000	1.000	BSI	RSI
85	0.000	0.000	1.000	BSI	RSI
86	0.199	0.000	0.801	BSI	RSI
87	0.186	0.000	0.814	BSI	RSI
88	0.212	0.387	0.401	BSI	MSI
89	0.000	0.000	1.000	BSI	RSI
90	0.000	0.608	0.392	RSI	MSI
91	0.000	0.807	0.193	RSI	MSI
92	0.414	0.000	0.586	BSI	RSI
93	0.000	0.401	0.599	BSI	RSI
94	0.000	0.192	0.808	BSI	RSI
95	0.000	0.000	1.000	BSI	RSI

## 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 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.

Tabel B.1 (Tabel lanjutan...)

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
96	0.000	0.000	1.000	BSI	RSI
97	0.618	0.000	0.382	MSI	RSI
98	0.000	0.601	0.399	RSI	MSI
99	0.180	0.000	0.820	BSI	RSI
100	0.000	0.619	0.381	RSI	MSI
101	0.200	0.000	0.800	BSI	RSI
102	0.000	0.000	1.000	BSI	RSI
103	0.000	0.190	0.810	BSI	BSI
104	0.000	0.371	0.629	BSI	RSI
105	0.000	0.440	0.560	BSI	RSI
106	0.000	0.194	0.806	BSI	MSI
107	0.000	0.617	0.383	RSI	RSI
108	0.000	0.192	0.808	BSI	MSI
109	0.000	0.258	0.742	BSI	RSI

HASIL PREDIKSI PERCOBAAN 3 KELAS K=10

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
1	0.099	0.100	0.800	BSI	RSI
2	0.202	0.404	0.394	MSI	RSI
3	0.191	0.194	0.615	BSI	RSI
4	0.394	0.197	0.409	BSI	BSI
5	0.096	0.203	0.702	BSI	RSI
6	0.197	0.204	0.599	BSI	RSI
7	0.191	0.106	0.704	BSI	RSI
8	0.000	0.507	0.493	MSI	MSI
9	0.077	0.152	0.771	BSI	RSI
10	0.299	0.284	0.418	BSI	RSI
11	0.116	0.312	0.572	BSI	RSI
12	0.095	0.421	0.484	BSI	MSI
13	0.204	0.392	0.404	BSI	RSI
14	0.000	0.192	0.808	BSI	MSI
15	0.190	0.299	0.510	BSI	MSI
16	0.101	0.392	0.507	BSI	RSI
17	0.099	0.198	0.703	BSI	RSI
18	0.000	0.000	1.000	BSI	MSI
19	0.097	0.193	0.710	BSI	RSI
20	0.000	0.495	0.505	BSI	BSI
21	0.200	0.214	0.587	BSI	MSI
22	0.090	0.000	0.910	BSI	RSI
23	0.000	0.212	0.788	BSI	MSI
24	0.000	0.188	0.812	BSI	MSI
25	0.000	0.000	1.000	BSI	MSI
26	0.208	0.096	0.695	BSI	MSI
27	0.097	0.509	0.394	MSI	MSI
28	0.698	0.000	0.302	RSI	BSI
29	0.000	0.000	1.000	BSI	RSI
30	0.293	0.488	0.219	MSI	MSI
31	0.105	0.000	0.895	BSI	RSI

Tabel B.2 (Tabel lanjutan...)

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
32	0.088	0.088	0.823	BSI	RSI
33	0.100	0.495	0.405	MSI	RSI
34	0.000	0.601	0.399	MSI	MSI
35	0.206	0.401	0.393	MSI	BSI
36	0.097	0.304	0.599	BSI	MSI
37	0.000	0.084	0.916	BSI	RSI
38	0.000	0.191	0.809	BSI	RSI
39	0.000	0.103	0.897	BSI	RSI
40	0.097	0.304	0.598	BSI	RSI
41	0.092	0.092	0.817	BSI	RSI
42	0.197	0.402	0.401	MSI	MSI
43	0.193	0.292	0.515	BSI	RSI
44	0.000	0.000	1.000	BSI	RSI
45	0.000	0.413	0.587	BSI	MSI
46	0.399	0.096	0.504	BSI	MSI
47	0.295	0.100	0.605	BSI	RSI
48	0.298	0.097	0.605	BSI	MSI
49	0.305	0.198	0.498	BSI	RSI
50	0.000	0.000	1.000	BSI	RSI
51	0.188	0.188	0.624	BSI	RSI
52	0.000	0.443	0.557	BSI	MSI
53	0.000	0.217	0.783	BSI	RSI
54	0.099	0.311	0.591	BSI	MSI
55	0.000	0.409	0.591	BSI	RSI
56	0.093	0.413	0.494	BSI	MSI
57	0.000	0.000	1.000	BSI	MSI
58	0.298	0.194	0.509	BSI	RSI
59	0.000	0.399	0.601	BSI	MSI
60	0.099	0.214	0.687	BSI	MSI
61	0.204	0.292	0.504	BSI	RSI
62	0.078	0.157	0.765	BSI	RSI
63	0.098	0.200	0.702	BSI	RSI
64	0.000	0.301	0.699	BSI	MSI
65	0.000	0.205	0.795	BSI	RSI
66	0.295	0.200	0.505	BSI	MSI
67	0.000	0.728	0.272	MSI	MSI
68	0.406	0.099	0.495	BSI	RSI
69	0.304	0.395	0.301	MSI	RSI
70	0.307	0.000	0.693	BSI	RSI
71	0.292	0.506	0.202	MSI	MSI
72	0.096	0.297	0.607	BSI	RSI
73	0.000	0.396	0.604	BSI	RSI
74	0.293	0.205	0.502	BSI	RSI
75	0.101	0.095	0.804	BSI	BSI
76	0.196	0.098	0.706	BSI	RSI
77	0.104	0.197	0.699	BSI	RSI
78	0.000	0.000	1.000	BSI	RSI
79	0.097	0.189	0.714	BSI	RSI
80	0.096	0.100	0.804	BSI	RSI
81	0.000	0.336	0.664	BSI	MSI

## 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 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.

Tabel B.2 (Tabel lanjutan...)

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
82	0.201	0.200	0.599	BSI	BSI
83	0.309	0.099	0.593	BSI	RSI
84	0.302	0.099	0.600	BSI	RSI
85	0.199	0.611	0.190	MSI	RSI
86	0.000	0.096	0.904	BSI	RSI
87	0.266	0.092	0.642	BSI	RSI
88	0.100	0.100	0.800	BSI	MSI
89	0.000	0.335	0.665	BSI	MSI
90	0.000	0.310	0.690	BSI	MSI
91	0.506	0.000	0.494	RSI	RSI
92	0.102	0.196	0.702	BSI	RSI
93	0.098	0.316	0.587	BSI	RSI
94	0.000	0.196	0.804	BSI	RSI
95	0.091	0.305	0.604	BSI	RSI
96	0.310	0.100	0.589	BSI	BSI
97	0.101	0.308	0.591	BSI	RSI
98	0.100	0.102	0.798	BSI	RSI
99	0.199	0.102	0.698	BSI	RSI
100	0.097	0.199	0.704	BSI	RSI
101	0.000	0.194	0.806	BSI	RSI
102	0.407	0.393	0.200	RSI	RSI
103	0.000	0.310	0.690	BSI	BSI
104	0.401	0.307	0.292	RSI	RSI
105	0.104	0.298	0.599	BSI	RSI
106	0.401	0.000	0.599	BSI	RSI
107	0.114	0.394	0.492	BSI	RSI
108	0.407	0.108	0.485	BSI	BSI
109	0.097	0.235	0.668	BSI	RSI

HASIL PREDIKSI PERCOBAAN 3 KELAS K=15

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
1	0.000	0.126	0.874	MSI	MSI
2	0.211	0.394	0.394	BSI	BSI
3	0.247	0.061	0.692	MSI	RSI
4	0.061	0.189	0.751	MSI	RSI
5	0.000	0.613	0.387	BSI	MSI
6	0.133	0.208	0.659	MSI	MSI
7	0.065	0.262	0.673	MSI	RSI
8	0.064	0.263	0.673	MSI	RSI
9	0.067	0.267	0.666	MSI	RSI
10	0.189	0.125	0.686	MSI	RSI
11	0.258	0.226	0.515	MSI	MSI
12	0.241	0.219	0.540	MSI	MSI
13	0.329	0.000	0.671	MSI	RSI
14	0.247	0.251	0.502	MSI	BSI
15	0.347	0.192	0.461	MSI	RSI
16	0.314	0.064	0.622	MSI	RSI
17	0.204	0.131	0.665	MSI	RSI

Tabel B.3 (Tabel lanjutan...)

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
18	0.067	0.260	0.674	MSI	RSI
19	0.133	0.326	0.541	MSI	RSI
20	0.130	0.259	0.612	MSI	BSI
21	0.210	0.050	0.740	MSI	RSI
22	0.000	0.289	0.711	MSI	RSI
23	0.000	0.127	0.873	MSI	RSI
24	0.000	0.394	0.606	MSI	RSI
25	0.274	0.335	0.391	MSI	RSI
26	0.129	0.202	0.668	MSI	MSI
27	0.131	0.127	0.741	MSI	RSI
28	0.401	0.067	0.532	MSI	RSI
29	0.129	0.127	0.744	MSI	RSI
30	0.000	0.381	0.619	MSI	RSI
31	0.196	0.138	0.666	MSI	RSI
32	0.058	0.276	0.666	MSI	RSI
33	0.129	0.481	0.390	BSI	RSI
34	0.132	0.392	0.476	MSI	RSI
35	0.198	0.267	0.536	MSI	BSI
36	0.000	0.122	0.878	MSI	RSI
37	0.134	0.399	0.467	MSI	MSI
38	0.474	0.198	0.328	RSI	BSI
39	0.132	0.412	0.457	MSI	RSI
40	0.063	0.139	0.798	MSI	RSI
41	0.000	0.000	1.000	MSI	RSI
42	0.130	0.258	0.612	MSI	RSI
43	0.076	0.274	0.650	MSI	RSI
44	0.131	0.204	0.665	MSI	RSI
45	0.341	0.074	0.585	MSI	BSI
46	0.000	0.196	0.804	MSI	MSI
47	0.068	0.454	0.478	MSI	RSI
48	0.065	0.463	0.472	MSI	RSI
49	0.136	0.198	0.667	MSI	RSI
50	0.263	0.203	0.534	MSI	RSI
51	0.126	0.130	0.744	MSI	RSI
52	0.000	0.131	0.869	MSI	RSI
53	0.066	0.342	0.592	MSI	RSI
54	0.125	0.383	0.492	MSI	MSI
55	0.211	0.262	0.527	MSI	BSI
56	0.200	0.403	0.397	BSI	RSI
57	0.000	0.406	0.594	MSI	RSI
58	0.000	0.195	0.805	MSI	RSI
59	0.000	0.628	0.372	BSI	MSI
60	0.210	0.266	0.524	MSI	BSI
61	0.216	0.314	0.470	MSI	BSI
62	0.134	0.401	0.465	MSI	RSI
63	0.193	0.236	0.571	MSI	MSI
64	0.132	0.654	0.214	BSI	MSI
65	0.325	0.193	0.482	MSI	RSI
66	0.064	0.428	0.508	MSI	MSI
67	0.138	0.198	0.663	MSI	RSI

## 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 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.

Tabel B.3 (Tabel lanjutan...)

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
68	0.000	0.000	1.000	MSI	BSI
69	0.275	0.198	0.527	MSI	RSI
70	0.064	0.340	0.595	MSI	MSI
71	0.190	0.388	0.422	MSI	RSI
72	0.199	0.129	0.671	MSI	RSI
73	0.125	0.348	0.527	MSI	RSI
74	0.000	0.264	0.736	MSI	RSI
75	0.128	0.208	0.663	MSI	RSI
76	0.116	0.352	0.532	MSI	RSI
77	0.258	0.000	0.742	MSI	RSI
78	0.273	0.327	0.399	MSI	BSI
79	0.196	0.336	0.469	MSI	RSI
80	0.063	0.392	0.545	MSI	RSI
81	0.000	1.000	0.000	BSI	RSI
82	0.066	0.161	0.773	MSI	RSI
83	0.204	0.196	0.601	MSI	RSI
84	0.260	0.261	0.479	MSI	RSI
85	0.127	0.126	0.747	MSI	RSI
86	0.321	0.066	0.613	MSI	RSI
87	0.000	1.000	0.000	BSI	MSI
88	0.272	0.190	0.538	MSI	RSI
89	0.068	0.201	0.731	MSI	RSI
90	0.000	0.191	0.809	MSI	RSI
91	0.206	0.000	0.794	MSI	BSI
92	0.000	0.470	0.530	MSI	MSI
93	0.197	0.278	0.525	MSI	MSI
94	0.065	0.400	0.535	MSI	MSI
95	0.000	0.200	0.800	MSI	RSI
96	0.000	0.000	1.000	MSI	MSI
97	0.068	0.325	0.607	MSI	RSI
98	0.132	0.334	0.533	MSI	RSI
99	0.268	0.267	0.465	MSI	MSI
100	0.064	0.000	0.936	MSI	RSI
101	0.064	0.414	0.522	MSI	MSI
102	0.132	0.130	0.738	MSI	RSI
103	0.285	0.260	0.454	MSI	RSI
104	0.000	0.567	0.433	BSI	MSI
105	0.131	0.410	0.459	MSI	RSI
106	0.000	0.132	0.868	MSI	RSI
107	0.060	0.325	0.615	MSI	MSI
108	0.203	0.000	0.797	MSI	RSI
109	0.201	0.332	0.467	MSI	RSI

HASIL PREDIKSI PERCOBAAN 3 KELAS K=20

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
1	0.049	0.146	0.805	MSI	RSI
2	0.100	0.294	0.606	MSI	RSI
3	0.162	0.444	0.394	BSI	RSI

Tabel B.4 (Tabel lanjutan...)

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
4	0.098	0.200	0.702	MSI	RSI
5	0.141	0.283	0.576	MSI	RSI
6	0.302	0.051	0.647	MSI	MSI
7	0.160	0.392	0.448	MSI	RSI
8	0.099	0.000	0.901	MSI	RSI
9	0.195	0.049	0.757	MSI	RSI
10	0.095	0.298	0.607	MSI	RSI
11	0.148	0.356	0.496	MSI	RSI
12	0.091	0.046	0.863	MSI	RSI
13	0.095	0.146	0.760	MSI	RSI
14	0.049	0.350	0.600	MSI	MSI
15	0.000	0.618	0.382	BSI	MSI
16	0.360	0.287	0.353	RSI	BSI
17	0.054	0.352	0.594	MSI	RSI
18	0.199	0.247	0.554	MSI	RSI
19	0.049	0.196	0.754	MSI	RSI
20	0.050	0.352	0.598	MSI	RSI
21	0.102	0.197	0.701	MSI	BSI
22	0.049	0.195	0.755	MSI	RSI
23	0.103	0.196	0.701	MSI	RSI
24	0.223	0.329	0.448	MSI	BSI
25	0.147	0.245	0.608	MSI	RSI
26	0.366	0.104	0.529	MSI	BSI
27	0.208	0.118	0.673	MSI	RSI
28	0.145	0.297	0.558	MSI	RSI
29	0.154	0.248	0.599	MSI	MSI
30	0.000	0.569	0.431	BSI	MSI
31	0.000	0.355	0.645	MSI	RSI
32	0.049	0.102	0.849	MSI	RSI
33	0.144	0.144	0.712	MSI	RSI
34	0.197	0.203	0.600	MSI	MSI
35	0.088	0.321	0.591	MSI	RSI
36	0.000	0.525	0.475	BSI	MSI
37	0.167	0.253	0.580	MSI	RSI
38	0.100	0.249	0.651	MSI	RSI
39	0.094	0.190	0.716	MSI	MSI
40	0.248	0.298	0.454	MSI	BSI
41	0.097	0.098	0.805	MSI	RSI
42	0.000	0.000	1.000	MSI	MSI
43	0.147	0.240	0.614	MSI	RSI
44	0.096	0.102	0.802	MSI	MSI
45	0.263	0.191	0.546	MSI	BSI
46	0.351	0.295	0.353	MSI	RSI
47	0.099	0.149	0.751	MSI	RSI
48	0.000	0.333	0.667	MSI	MSI
49	0.143	0.564	0.294	BSI	MSI
50	0.051	0.352	0.597	MSI	MSI
51	0.100	0.148	0.753	MSI	RSI
52	0.193	0.142	0.666	MSI	RSI
53	0.000	0.334	0.666	MSI	RSI

## 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.



Tabel B.4 (Tabel lanjutan...)

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
54	0.049	0.097	0.854	MSI	RSI
55	0.049	0.451	0.500	MSI	MSI
56	0.257	0.296	0.447	MSI	BSI
57	0.154	0.346	0.500	MSI	RSI
58	0.000	0.440	0.560	MSI	RSI
59	0.144	0.191	0.665	MSI	MSI
60	0.151	0.196	0.653	MSI	BSI
61	0.146	0.199	0.656	MSI	RSI
62	0.142	0.417	0.441	MSI	MSI
63	0.000	0.349	0.651	MSI	RSI
64	0.047	0.143	0.810	MSI	RSI
65	0.049	0.241	0.710	MSI	RSI
66	0.148	0.300	0.552	MSI	RSI
67	0.316	0.145	0.539	MSI	RSI
68	0.266	0.342	0.392	MSI	RSI
69	0.000	0.000	1.000	MSI	MSI
70	0.148	0.408	0.445	MSI	RSI
71	0.148	0.197	0.655	MSI	RSI
72	0.138	0.453	0.409	BSI	RSI
73	0.000	0.193	0.807	MSI	RSI
74	0.000	0.144	0.856	MSI	RSI
75	0.047	0.400	0.552	MSI	RSI
76	0.102	0.098	0.800	MSI	RSI
77	0.265	0.297	0.438	MSI	MSI
78	0.110	0.247	0.643	MSI	MSI
79	0.099	0.243	0.658	MSI	MSI
80	0.244	0.211	0.545	MSI	RSI
81	0.051	0.253	0.695	MSI	MSI
82	0.197	0.150	0.653	MSI	MSI
83	0.149	0.349	0.502	MSI	MSI
84	0.000	0.343	0.657	MSI	RSI
85	0.193	0.366	0.441	MSI	MSI
86	0.000	1.000	0.000	BSI	RSI
87	0.155	0.291	0.554	MSI	RSI
88	0.094	0.563	0.343	BSI	MSI
89	0.205	0.249	0.546	MSI	BSI
90	0.202	0.297	0.501	MSI	RSI
91	0.098	0.209	0.693	MSI	RSI
92	0.049	0.296	0.655	MSI	MSI
93	0.149	0.195	0.656	MSI	BSI
94	0.235	0.189	0.575	MSI	RSI
95	0.000	0.296	0.704	MSI	RSI
96	0.191	0.339	0.471	MSI	MSI
97	0.147	0.295	0.559	MSI	RSI
98	0.199	0.150	0.650	MSI	BSI
99	0.096	0.148	0.756	MSI	RSI
100	0.048	0.146	0.806	MSI	RSI
101	0.050	0.293	0.657	MSI	RSI
102	0.093	0.485	0.423	BSI	MSI
103	0.100	0.453	0.446	BSI	MSI

## 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 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.

Tabel B.4 (Tabel lanjutan...)

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
104	0.000	0.394	0.606	MSI	RSI
105	0.050	0.147	0.803	MSI	BSI
106	0.245	0.445	0.310	BSI	RSI
107	0.189	0.096	0.716	MSI	MSI
108	0.198	0.244	0.559	MSI	RSI
109	0.112	0.099	0.790	MSI	BSI

HASIL PREDIKSI PERCOBAAN 3 KELAS K=25

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
1	0.081	0.446	0.473	BSI	MSI
2	0.037	0.311	0.652	BSI	RSI
3	0.158	0.195	0.647	BSI	RSI
4	0.241	0.234	0.525	BSI	RSI
5	0.123	0.402	0.475	BSI	RSI
6	0.157	0.243	0.601	BSI	MSI
7	0.080	0.436	0.485	BSI	MSI
8	0.079	0.000	0.921	BSI	RSI
9	0.114	0.314	0.572	BSI	RSI
10	0.202	0.274	0.523	BSI	RSI
11	0.040	0.312	0.648	BSI	MSI
12	0.038	0.079	0.883	BSI	RSI
13	0.234	0.121	0.645	BSI	RSI
14	0.078	0.078	0.844	BSI	RSI
15	0.233	0.216	0.551	BSI	MSI
16	0.077	0.306	0.617	BSI	RSI
17	0.116	0.272	0.612	BSI	MSI
18	0.000	0.000	1.000	BSI	MSI
19	0.154	0.124	0.722	BSI	RSI
20	0.199	0.197	0.604	BSI	BSI
21	0.116	0.196	0.688	BSI	RSI
22	0.083	0.526	0.390	MSI	MSI
23	0.204	0.239	0.557	BSI	BSI
24	0.154	0.243	0.603	BSI	RSI
25	0.205	0.082	0.714	BSI	RSI
26	0.297	0.195	0.509	BSI	BSI
27	0.078	0.156	0.765	BSI	RSI
28	0.126	0.193	0.681	BSI	MSI
29	0.120	0.282	0.598	BSI	BSI
30	0.120	0.364	0.516	BSI	RSI
31	0.272	0.292	0.436	BSI	MSI
32	0.196	0.278	0.526	BSI	RSI
33	0.199	0.279	0.522	BSI	MSI
34	0.121	0.197	0.682	BSI	BSI
35	0.118	0.276	0.606	BSI	RSI
36	0.080	0.127	0.793	BSI	RSI
37	0.326	0.311	0.363	BSI	RSI
38	0.107	0.143	0.749	BSI	RSI
39	0.161	0.156	0.683	BSI	RSI

Tabel B.5 (Tabel lanjutan...)

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
40	0.251	0.236	0.513	BSI	RSI
41	0.260	0.341	0.399	BSI	RSI
42	0.116	0.234	0.650	BSI	BSI
43	0.080	0.320	0.600	BSI	RSI
44	0.118	0.285	0.597	BSI	MSI
45	0.000	0.326	0.674	BSI	RSI
46	0.114	0.247	0.639	BSI	MSI
47	0.000	1.000	0.000	MSI	MSI
48	0.119	0.275	0.606	BSI	RSI
49	0.077	0.446	0.477	BSI	MSI
50	0.226	0.345	0.429	BSI	BSI
51	0.040	0.198	0.762	BSI	MSI
52	0.120	0.244	0.636	BSI	RSI
53	0.200	0.157	0.643	BSI	MSI
54	0.164	0.322	0.514	BSI	BSI
55	0.081	0.403	0.517	BSI	RSI
56	0.238	0.118	0.644	BSI	RSI
57	0.115	0.231	0.655	BSI	MSI
58	0.125	0.157	0.718	BSI	MSI
59	0.076	0.116	0.808	BSI	RSI
60	0.000	0.230	0.770	BSI	MSI
61	0.077	0.523	0.400	MSI	MSI
62	0.042	0.322	0.637	BSI	BSI
63	0.199	0.268	0.534	BSI	RSI
64	0.135	0.281	0.584	BSI	RSI
65	0.076	0.233	0.691	BSI	RSI
66	0.080	0.281	0.639	BSI	RSI
67	0.038	0.360	0.601	BSI	RSI
68	0.040	0.280	0.680	BSI	MSI
69	0.124	0.155	0.721	BSI	BSI
70	0.155	0.077	0.768	BSI	RSI
71	0.210	0.239	0.551	BSI	RSI
72	0.039	0.320	0.642	BSI	RSI
73	0.200	0.040	0.759	BSI	RSI
74	0.241	0.159	0.600	BSI	BSI
75	0.118	0.116	0.766	BSI	RSI
76	0.242	0.159	0.599	BSI	RSI
77	0.120	0.364	0.516	BSI	RSI
78	0.038	0.274	0.688	BSI	RSI
79	0.135	0.275	0.590	BSI	BSI
80	0.238	0.364	0.398	BSI	RSI
81	0.231	0.202	0.567	BSI	RSI
82	0.000	0.264	0.736	BSI	RSI
83	0.194	0.117	0.689	BSI	RSI
84	0.157	0.078	0.765	BSI	RSI
85	0.157	0.158	0.685	BSI	RSI
86	0.080	0.249	0.671	BSI	MSI
87	0.079	0.199	0.722	BSI	RSI
88	0.079	0.279	0.642	BSI	RSI
89	0.078	0.116	0.805	BSI	RSI

## 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.



Tabel B.5 (Tabel lanjutan...)

No	Probabilitas			Kelas	
	MSI	RSI	BSI	Prediksi	Aktual
90	0.191	0.231	0.578	BSI	RSI
91	0.151	0.379	0.470	BSI	MSI
92	0.118	0.323	0.559	BSI	MSI
93	0.080	0.287	0.633	BSI	BSI
94	0.137	0.314	0.549	BSI	BSI
95	0.203	0.200	0.597	BSI	RSI
96	0.126	0.114	0.760	BSI	RSI
97	0.078	0.357	0.566	BSI	RSI
98	0.151	0.291	0.559	BSI	MSI
99	0.088	0.357	0.554	BSI	RSI
100	0.077	0.115	0.808	BSI	MSI
101	0.198	0.439	0.363	MSI	RSI
102	0.199	0.120	0.681	BSI	RSI
103	0.282	0.198	0.520	BSI	MSI
104	0.243	0.196	0.561	BSI	RSI
105	0.077	0.326	0.597	BSI	RSI
106	0.116	0.326	0.557	BSI	BSI
107	0.000	0.455	0.545	BSI	MSI
108	0.077	0.160	0.763	BSI	RSI
109	0.157	0.207	0.636	BSI	MSI



**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.



## DAFTAR RIWAYAT HIDUP

Penulis bernama Yollla Damara ini dilahirkan di Kota Pariaman, Sumatra Barat, pada tanggal 31 Maret 1998. Penulis merupakan puteri pertama dari Ayahanda Damrah dan Ibunda Nelmawati.

Pendidikan yang ditempuh penulis dimulai dari Taman Kanak-Kanak (TK) di TK Diniyah Putri pada tahun 2003-2004, melanjutkan pendidikan dasar di SDN 030 Pekanbaru pada tahun 2004-2010, dilanjutkan dengan pendidikan menengah di SMPN 20 Pekanbaru pada tahun 2010-2013 dan SMKS Kansai

Pekanbaru pada 2013-2016, kemudian melanjutkan pendidikan Strata Satu (S1) dengan mengambil Program Studi Sistem Informasi di Perguruan Tinggi Negeri UIN Sultan Syarif Kasim Riau yang terletak di Kota Pekanbaru, Provinsi Riau.

Penelitian tugas akhir berjudul "**KLASIFIKASI DOKUMEN SCIENTIFIC PAPER GOOGLE SCHOLAR DOSEN SISTEM INFORMASI DENGAN METODE K-NEAREST NEIGHBOR**". Untuk menjalin komunikasi dengan penulis dapat menghubungi melalui e-mail [yolla.damara@uin-suska.ac.id](mailto:yolla.damara@uin-suska.ac.id).



© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim 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.



UIN SUSKA RIAU

UIN SUSKA RIAU