

PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK  
MENGELOMPOKKAN DATA OPINI PENGGUNA TWITTER MENGENAI  
PANDEMI COVID-19

Tugas Akhir disusun untuk memenuhi syarat  
mencapai gelar Kesarjanaan Komputer pada  
Program Studi Sistem Informasi  
Jenjang Program Strata-1



oleh:

**DIKOGIGIH PRASETYANTORO**

17.01.55.0073

22412

**FAKULTAS TEKNOLOGI INFORMASI  
UNIVERSITAS STIKUBANK (UNISBANK)**

SEMARANG

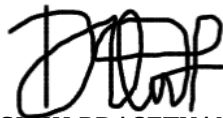
2021

## PERNYATAAN KESIAPAN UJIAN AKHIR

Saya DIKOGIGIH PRASETYANTORO, dengan ini menyatakan bahwa laporan Tugas Akhir yang berjudul :

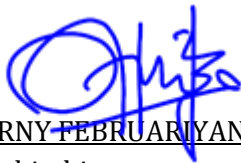
PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK  
MENGELOMPOKKAN DATA OPINI PENGGUNA TWITTER  
MENGENAI PANDEMI COVID-19.

adalah benar hasil karya saya dan belum pernah diajukan sebagai karya ilmiah sebagian atau seluruhnya atau pihak lain.



DIKOGIGIH PRASETYANTORO  
17.01.55.0073

Disetujui Oleh Pembimbing  
Kami setuju laporan tersebut diajukan untuk Ujian Tugas  
Akhir Semarang, 06 Juli 2021



HERNY FEBRUARIYANTI, S.T., M.Cs.  
Pembimbing



Dokumen ini diterbitkan secara elektronik. Disertai QRCode untuk validasi.

## SURAT PERNYATAAN KEASLIAN TUGAS AKHIR/ SKRIPSI

Yang bertanda tangan di bawah ini, saya menyatakan bahwa TUGAS AKHIR / SKRIPSI dengan Judul :

**PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK MENGELOMPOKKAN DATA OPINI PENGGUNA TWITTER MENGENAI PANDEMI COVID-19.**

yang telah diuji di depan tim penguji pada tanggal 13-07-2021, adalah benar hasil karya saya dan dalam TUGAS AKHIR /SKRIPSI ini tidak terdapat keseluruhan atau sebagian tulisan orang lain yang saya ambil dengan cara menyalin, atau meniru dalam bentuk rangkaian kalimat atau simbol yang saya aku seolah-olah sebagai tulisan saya sendiri dan atau tidak terdapat bagian atau keseluruhan tulisan yang saya salin, tiru atau yang saya ambil dari tulisan orang lain tanpa memberikan pengakuan pada penulis aslinya.

Apabila saya melakukan hal tersebut diatas, baik sengaja maupun tidak, dengan ini saya menyatakan menarik TUGAS AKHIR / SKRIPSI yang saya ajukan sebagai hasil tulisan saya sendiri.

Bila kemudian terbukti bahwa saya ternyata melakukan tindakan menyalin atau meniru tulisan orang lain seolah-olah hasil pemikiran saya sendiri, berarti gelar dan ijazah saya yang telah diberikan oleh Universitas Sikubank (UNISBANK) Semarang batal saya terima.

Semarang, 13-07-2021  
Yang Menyatakan  
  
(DIKOGIGIH PRASETYANTORO)  
NIM. 17.01.55.0073

SAKSI 1  
Tim Penguji



(HERNY FEBRUARIYANTI, ST, M.Cs.)

SAKSI 2  
Tim Penguji



(ARIEF JANANTO, S.Kom., M.Cs.)

SAKSI 3  
Tim Penguji



(SULASTRI, DRA., M.KOM)



Document ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan Dinas Unisbank. Untuk memastikan keasliannya, silahkan scan QRCode dan pastikan diarahkan ke alamat <https://digsig.unisbank.ac.id>

1/1

---

## HALAMAN PENGESAHAN

Tugas Akhir dengan Judul  
**PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK MENGELOMPOKKAN  
DATA OPINI PENGGUNA TWITTER MENGENAI PANDEMI COVID-19.**

Ditulis oleh  
NIM : 17.01.55.0073  
Nama : DIKOGIGIH PRASETYANTORO

Telah dipertahankan di depan Tim Dosen Penguji Tugas Akhir dan diterima sebagai salah satu syarat  
guna menyelesaikan Jenjang Program S1 Program Studi pada Fakultas Universitas Stikubank  
(UNISBANK) Semarang.

Semarang, 02-08-2021  
Ketua



(HERNY FEBRUARIYANTI, ST, M.Cs)  
NIDN. 0614027301

Sekretaris



(ARIEF JANANTO, S.Kom., M.Cs.)  
NIDN. 0606017402

Anggota



(SULASTRI, DRA, M.KOM)  
NIDN. 0608126601

Mengetahui,  
Universitas Stikubank (UNISBANK) Semarang  
Fakultas Teknologi Informasi  
Dekan



(KRISTOPHORUS HADIONO, Ph.D)  
NIDN. 0622027601



## **MOTTO DAN PERSEMBAHAN**

Motto

“AKU TIDAK BERKOMENTAR TENTANG CARA ORANG MENJALANI  
HIDUP”

- Patrick Star

Persembahan

Skripsi ini saya persembahkan untuk:

1. Kedua orang tua sebagai tanda terima kasih yang selalu mendoakan dan mengorbankan segalanya untuk masa depan anaknya.
2. Para sahabat dan teman-teman yang selalu membantu.
3. Ibu Herny Februriyanti, S.T., M.Cs selaku dosen pembimbing.
4. Para dosen Program Studi Sistem Informasi yang senantiasa membagikan ilmunya.

## KATA PENGANTAR

*Assalamualaikum Wr.Wb.*

Dengan memanjatkan Puji syukur serta kehadiran Allah SWT, atas rahmat taufiq dan hidayah-Nya yang telah dilimpahkan, sehingga pada kesempatan ini penulis dapat menyelesaikan penelitian tugas akhir dengan judul “Penerapan Algoritma K-Means Clustering Untuk Mengelompokkan Data Opini Pengguna Twitter Mengenai Pandemi Covid-19”.

Didalam penyelesaian penelitian tugas akhir ini, ada berbagai pihak yang memberikan bantuan dan memberikan dukungan semangat kepada penulis. Oleh karena itu penulis menyampaikan doa serta ucapan terima kasih kepada:

1. Dr. Safik Faozi, SH., M.Hum selaku Rektor Universitas Stikubank (UNISBANK) Semarang.
2. Kristophorus Hadiono, S.Kom., Ph.D selaku Dekan Fakultas Teknologi Informasi Universitas Stikubank (UNISBANK) Semarang.
3. Arief Jananto, S.Kom., M.Cs selaku ketua Progdi jurusan Sistem Informasi Universitas Stikubank (UNISBANK) Semarang.
4. Hery Februariyanti, ST, M. Cs selaku Dosen Pembimbing Tugas Akhir yang dengan sabar membimbing, mengarahkan, memberi masukan, serta meluangkan waktunya kepada penulis dalam penyelesaian pengerjaan tugas akhir ini.
5. Seluruh Dosen, dan Staff Administrasi Universitas Stikubank yang telah banyak membantu dan memberikan kemudahan, terima kasih atas segala ilmu dan bimbingannya.
6. Kedua orang tua penulis yang telah memberikan dukungan dan semangat agar dapat mengerjakan tugas akhir dengan baik dan lancar.
7. Seluruh sahabat dan teman-teman studi dari penulis yang tidak dapat disebut satu persatu, yang juga telah memberikan semangat dukungan dan kebaikan kepada penulis.

Penulis menyadari bahwa masih terdapat beberapa kesalahan dalam penulisan tugas akhir, oleh karena itu penulis sangat mengharap segala kritik dan saran yang bersifat membangun terhadap penulisan tugas akhir ini. Penulis berharap penelitian tugas akhir ini dapat memberikan manfaat bagi penulis secara pribadi dan bagi pembaca nantinya.

*Wassalamualaikum Wr.Wb.*

Semarang, Juli 2021

Penulis

## DAFTAR ISI

PERNYATAAN KESIAPAN UJIAN AKHIR.....	ii
SURAT PERNYATAAN KEASLIAN TUGAS AKHIR/ SKRIPSI .....	iii
HALAMAN PENGESAHAN.....	iv
MOTTO DAN PERSEMBAHAN .....	v
ABSTRAK .....	vi
KATA PENGANTAR .....	viii
DAFTAR ISI.....	xiii
DAFTAR TABEL.....	xii
DAFTAR GAMBAR .....	xiii
BAB I.....	1
PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Perumusan Masalah.....	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penulisan.....	4
BAB II.....	6
TINJAUAN PUSTAKA .....	6
2.1 Penelitian Terdahulu.....	6
2.2 Perbedaan dengan Penelitian Terdahulu .....	11
2.3 Twitter .....	12
2.4 <i>Text Mining</i> .....	13

2.5	<i>Information Retrieval (IR)</i> .....	14
2.6	<i>Text Preprocessing</i> .....	15
2.7	TF-IDF ( <i>Term Frequency-Inverse Document Frequency</i> ).....	16
2.8	<i>Davies Bouldin Index (DBI)</i> .....	17
2.9	<i>Clustering</i> .....	18
2.10	Algoritma <i>K-Means</i> .....	19
2.11	Visualisasi.....	21
BAB III .....		22
METODE PENELITIAN.....		22
3.1	Pengumpulan Data .....	23
3.2	<i>Text Preprocessing</i> .....	27
3.3	TF-IDF ( <i>Term Frequency-Inverse Dense Frequency</i> ) .....	33
3.4	<i>Clustering</i> .....	37
3.4.1.	<i>Davies Bouldin Index (DBI)</i> .....	37
3.4.2.	Implementasi Algoritma <i>K-Means</i> .....	39
BAB IV .....		42
HASIL PENELITIAN DAN PEMBAHASAN .....		42
4.1	Hasil Penelitian.....	42
4.1.1	Instalasi <i>Package</i> .....	42
4.1.2	Memanggil <i>Package</i> .....	44
4.1.3	Membaca Data .....	45
4.1.4	<i>Text Preprocessing</i> .....	47
4.1.5	TF-IDF ( <i>Term Frequency-Inverse Dense Frequency</i> ).....	54
4.1.6	<i>K-Means Clustering</i> .....	57
4.1.7	Visualisasi Data.....	58



4.2 Pembahasan .....	61
BAB V.....	63
KESIMPULAN DAN SARAN.....	63
5.1 Kesimpulan.....	63
5.2 Saran.....	64
DAFTAR PUSTAKA .....	65
LAMPIRAN.....	68
LISTING PROGRAM .....	68
LAMPIRAN.....	78
LEMBAR BIMBINGAN .....	78
LAMPIRAN.....	80
SURAT KETERANGAN BEBAS PLAGIARISME TURNITIN.....	80
LAMPIRAN.....	84
BERITA ACARA REVISI UJIAN TUGAS AKHIR/ SKRIPSI.....	84