

**ANALISIS SENTIMEN MENGGUNAKAN K-NEAREST NEIGHBOR  
TERHADAP *INITIAL PUBLIC OFFERING* (IPO) PT GOTO GOJEK  
TOKOPEDIA TBK PADA TWITTER**

**SKRIPSI**



**Disusun oleh**

**Irvan Muhammad Ramadhan**

**16.11.0822**

**PROGRAM SARJANA**

**PROGRAM STUDI INFORMATIKA**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS AMIKOM YOGYAKARTA**

**2023**

**ANALISIS SENTIMEN MENGGUNAKAN K-NEAREST NEIGHBOR  
TERHADAP INITIAL PUBLIC OFFERING (IPO) PT GOTO GOJEK  
TOKOPEDIA TBK PADA TWITTER**

**SKRIPSI**

untuk memenuhi sebagian  
persyaratan mencapai gelar  
Sarjana  
pada Program Studi Informatika



**disusun oleh**

**Irvan Muhammad Ramadhan**

**16.11.0822**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2023**

**ANALISIS SENTIMEN MENGGUNAKAN K-NEAREST  
NEIGHBOR TERHADAP *INITIAL PUBLIC OFFERING* (IPO)  
PT GOTO GOJEK TOKOPEDIA TBK PADA TWITTER**  
yang dipersiapkan dan disusun oleh  
**Irvan Muhammad Ramadhan**  
**16.11.0822**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 18 November 2022

**Dosen Pembimbing**



**Erni Seniwati, S.Kom, M.Cs**

**NIK. 190302231**

# PENGESAHAN

## SKRIPSI

### ANALISIS SENTIMEN MENGGUNAKAN K-NEAREST NEIGHBOR TERHADAP *INITIAL PUBLIC OFFERING (IPO) PT GOTO GOJEK* TOKOPEDIA TBK PADA TWITTER

yang dipersiapkan dan disusun oleh

**Irvan Muhammad Ramadhan**

**16.11.0822**

telah dipertahankan di depan Dewan Penguji

pada tanggal 18 November 2022


**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Erni Seniwati, S.Kom,M.CS.**

**NIK. 190302231**



**Ikamah, M.Kom.**

**NIK. 190302282**



**Rumini, M.Kom.**

**NIK. 190302246**

Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 27 Maret 2023

**DEKAN FAKULTAS ILMU KOMPUTER**



**Hanif Al Fatta, M.Kom.**

**NIK. 190302096**

## PERNYATAAN

Saya yang bertandatangan di bawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 16 Mei 2023



Irvan Muhammad Ramadhan

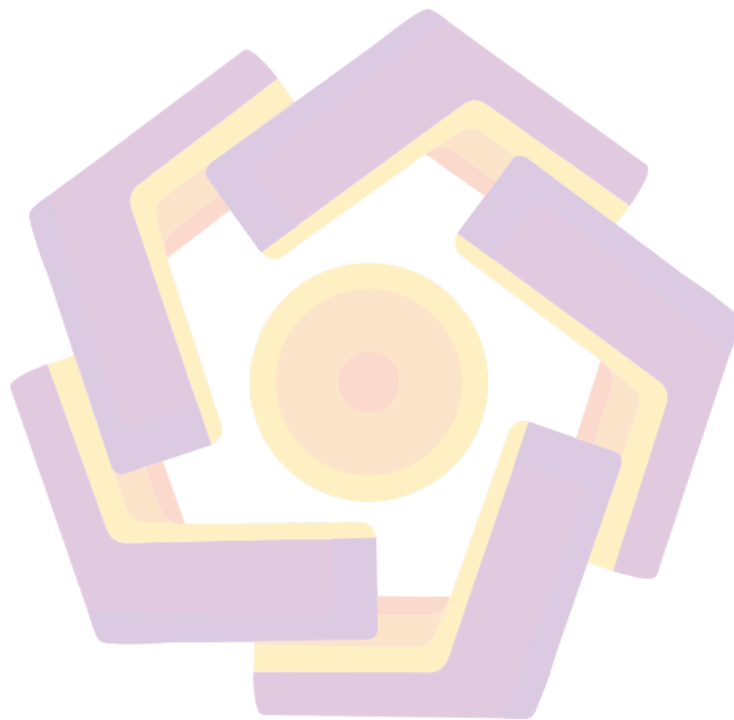
NIM 16.11.0822

## MOTTO

*”Hidup adalah tentang proses”*

*“Jalani apa yang ada di depan mata, jangan sampai terjebak di antara masa lalu dan masa depan sehingga melupakan sesuatu yang penting yaitu saat ini”*

*“Selesaikan apa yang sudah kamu pilih dan mulai”*



## PERSEMBAHAN

Bismillahirrahmanirrahim, Alhamdulillahrabbi'alamini puji syukur penulis panjatkan kepada Allah SWT atas Rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul “Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap *Initial Public Offering* (IPO) Pt Goto Gojek Tokopedia Tbk Pada Twitter”. Untuk memenuhi syarat akademis dalam menyelesaikan Program Studi Strata Satu (S1) pada Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.

Mengingat keterbatasan pengetahuan dan pengalaman dalam penulisan skripsi ini, Penulis banyak mendapatkan bimbingan, petunjuk, saran dan arahan dari berbagai pihak, oleh karena itu dengan kerendahan hati dan rasa hormat Penulis mengucapkan terimakasih sebesar-besarnya kepada :

1. Bapak Prof. Dr. M. Suyanto, MM. Selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Sudarmawan, S.T., M.T. Selaku Dosen wali yang telah menjadi wali saya selama menjadi mahasiswa.
3. Ibu Erni Seniwati, S.Kom., M.Cs. Selaku Dosen pembimbing yang telah membantu dalam penulisan skripsi ini.
4. Ibu Ikmah, M.Kom. dan Ibu Rumini, M.Kom. Selaku Dosen Penguji yang telah memberikan petunjuk, serta nasehat dalam ujian skripsi ini.
5. Kedua orang tua saya, Bapak Imam dan Ibu Surtiyah yang selalu memberikan do'a dan support yang tak tergantikan kepada penulis dalam pengerjaan skripsi ini.
6. Adik saya, Nadia yang selalu memberikan dukungan untuk masnya.
7. Teman – teman Keluarga Berwacana terlebih untuk Rista dan Andito yang telah membantu dan memberikan semangat serta solusi dalam pembuatan skripsi.
8. Teman – teman Teh Manis Kos sebagai tempat sharing dalam kelancaran progress skripsi penulis.
9. Teman – teman kelas S1-Informatika 13 (BCD Class) yang tidak dapat saya sebutkan satu persatu, terimakasih semuanya.
10. Teman - teman FOSSIL terlebih lagi untuk Tama dan Rama yang selalu mendukung dan memberikan arahan dalam pengerjaan.

## KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan rahmat serta hidayah-Nya dan shalawat serta salam juga tidak lupa penulis panjatkan kepada junjungan kita Nabi Muhammad SAW yang telah memberikan teladan mulia dalam menuntun ummatnya sehingga penulis dapat menyelesaikan skripsi ini.

Skripsi yang berjudul **"Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap Initial Public Offering (IPO) Pt Goto Gojek Tokopedia Tbk Pada Twitter"** ini disusun sebagai salah satu syarat utama untuk menyelesaikan program sarjana pada Universitas AMIKOM Yogyakarta.

Penyelesaian skripsi ini juga tidak lepas dari bantuan berbagai pihak, karena itu pada kesempatan ini penulis ingin menyampaikan rasa hormat dan terima kasih kepada :

1. Bapak Prof. Dr. M. Suyanto, MM. Selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Hanif Al Fatta, M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Ibu Erni Seniwati, S.Kom., M.Cs. selaku dosen pembimbing yang selalu bijaksana memberikan bimbingan, nasehat serta waktunya selama penulisan skripsi ini.
4. Ibu Ikmah, M.Kom. dan Ibu Rumini, M.kom sebagai dosen penguji serta semua dosen Prodi Informatika Universitas Amikom Yogyakarta, terima kasih atas semua jasa Bapak dan Ibu Dosen.
5. Kedua orang tua saya, dan adik tersayang yang telah mendo'akan, mendukung dan memberikan semangat.
6. Segenap Dosen dan Civitas Akademika Universitas AMIKOM Yogyakarta yang telah memberikan banyak ilmu dan pengalaman kepada penulis selama menjalani perkuliahan.
7. Teman-teman kelas Informatika 13 yang tidak dapat saya sebutkan satu persatu, terimakasih semuanya.
8. Seluruh teman-teman dan semua pihak yang tidak dapat saya sebutkan satu persatu, terimakasih banyak atas segala bantuannya dalam menyelesaikan karya ini.

Penulis menyadari skripsi ini masih ada kekurangan, maka dari itu kritik dan saran yang membangun serta teguran dari semua pihak, penulis menerima dengan lapang dada untuk kesempurnaan karya selanjutnya. Akhirnya kepada Allah SWT jualah tangan bertengadah dan berharap serta, semoga skripsi sederhana ini bermanfaat. Khususnya bagi penulis dan pembaca yang budiman pada umumnya. Apabila terdapat kesalahan semoga Allah melimpahkan maghfirah-Nya. *Aamiin yaa Kholiq.*

Yogyakarta, 28 Maret 2023

Irvan Muhammad Ramadhan  
NIM : 16.11.0822

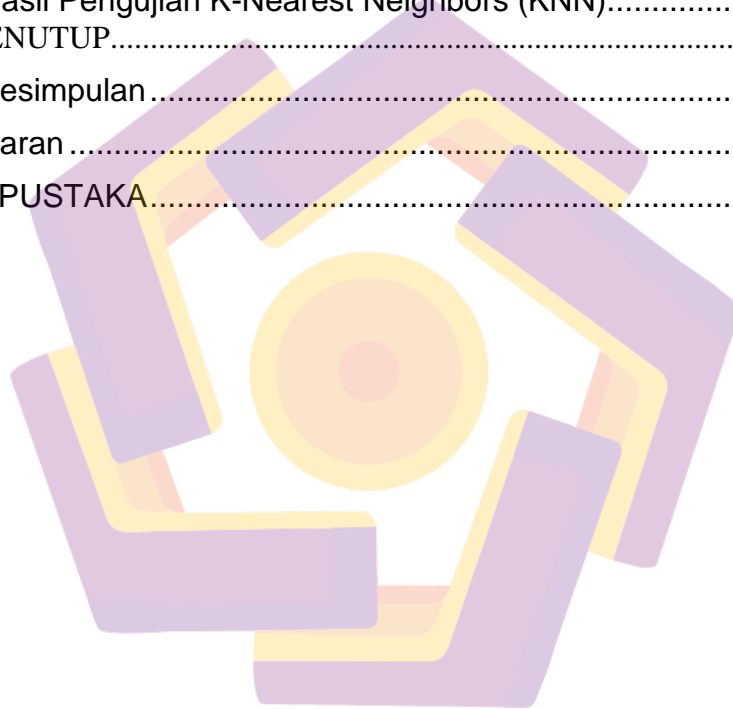


## DAFTAR ISI

ANALISIS SENTIMEN MENGGUNAKAN K-NEAREST NEIGHBOR TERHADAP <i>INITIAL PUBLIC OFFERING (IPO) PT GOTO GOJEK TOKOPEDIA TBK PADA TWITTER</i> .....	i
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL.....	xiii
INTISARI.....	xiv
<i>ABSTRACT</i> .....	xv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Masalah.....	2
1.4. Maksud dan Tujuan Penelitian.....	3
1.4.1. Maksud Penelitian.....	3
1.4.2. Tujuan Penelitian.....	3
1.5. Manfaat Penelitian.....	3
1.6. Metodologi Penelitian.....	4
1.6.1. Metode Pengumpulan Data.....	4
1.6.2. Metode Analisis.....	4
1.6.3. Metode Perancangan.....	5
1.6.4. Metode Pengujian.....	5
1.7. Sistematika Penelitian.....	5
BAB I PENDAHULUAN.....	5
BAB II LANDASAN TEORI.....	6
BAB III METODE PENELITIAN.....	6
BAB IV HASIL DAN PEMBAHASAN.....	6
BAB V PENUTUP.....	6
DAFTAR PUSTAKA.....	6
BAB II LANDASAN TEORI.....	7
2.1. Kajian Pustaka.....	7
2.2. Dasar Teori.....	8

2.2.1.	Data Mining.....	8
2.2.2.	Pengelompokan Data Mining.....	9
2.2.3.	Analisis Sentiment .....	10
2.2.4.	Twitter .....	11
2.2.5.	Analisis Sentimen Twitter .....	12
2.2.6.	Twitter API .....	12
2.2.7.	Rapidminer .....	13
2.2.8.	KNN (K-Nearest Neighbors) .....	13
2.2.9.	Text Mining .....	14
2.2.10.	Stemming .....	15
BAB III METODE PENELITIAN .....		16
3.1.	Analisis Masalah.....	17
3.2.	Merumuskan Masalah .....	17
3.3.	Penentuan Tujuan .....	17
3.4.	Studi Literatur .....	17
3.5.	Analisis dan Perancangan .....	18
3.5.1.	Analisis Kebutuhan Sistem .....	18
3.5.1.2.	Analisis Kebutuhan Non-Fungsional .....	18
3.5.2.	Analisis Data.....	19
3.5.3.	Penentuan Query.....	20
3.5.4.	Preprocessing.....	21
3.5.4.2	Case Folding.....	25
3.5.4.3	Tokenizing.....	27
3.5.4.4	Pelabelan Data.....	29
3.5.4.5	Filtering dan Stemming .....	29
3.5.5.	Analisis algoritma K-NN .....	31
3.5.6.	Perancangan system .....	32
3.5.7.	Diagram Alir Proses Data .....	33
3.5.8.	Pengujian Akurasi.....	33
BAB IV .....		35
HASIL DAN PEMBAHASAN.....		35
4.1.	Implementasi .....	35

4.1.1.	Data .....	35
4.1.2.	Preprocessing .....	36
4.1.2.1.	Cleaning .....	36
4.1.2.2.	Case Folding .....	38
4.1.2.3.	Tokenizing .....	38
4.1.2.4.	Filtering .....	39
4.1.2.5.	Stemming .....	40
4.2.	Hasil dan pembahasan .....	41
4.2.1.	Klasifikasi Menggunakan KNN .....	41
4.2.2.	Hasil Pengujian K-Nearest Neighbors (KNN) .....	42
BAB V PENUTUP .....		49
5.1.	Kesimpulan .....	49
5.2.	Saran .....	49
DAFTAR PUSTAKA .....		50

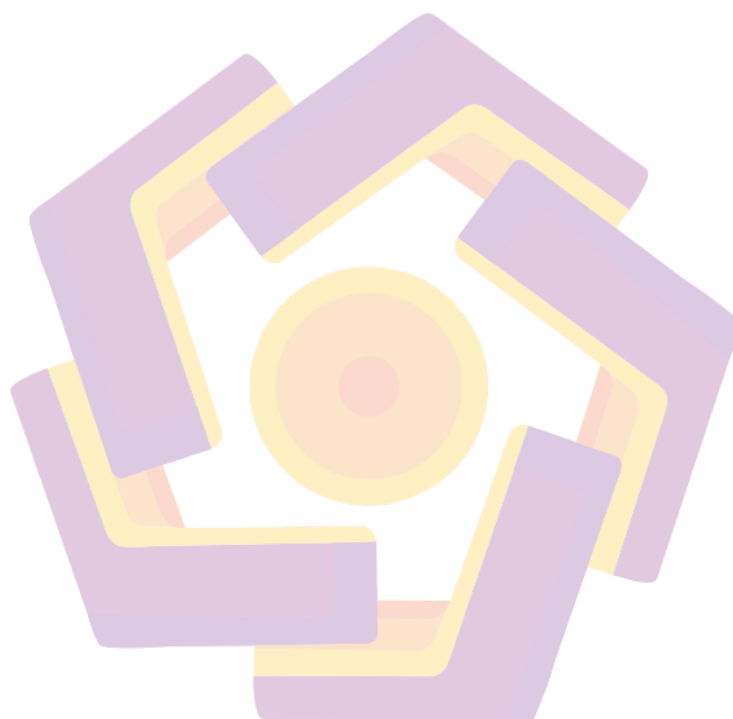


## DAFTAR GAMBAR

Gambar 3. 1 Alur Metode Penelitian.....	16
Gambar 3. 2 Diagram Analisis Data .....	20
Gambar 3. 3 Flowchart Penentuan Query.....	20
Gambar 3. 4 Proses Cleaning .....	21
Gambar 3. 5 Proses Case Folding .....	25
Gambar 3. 6 Proses Tokenizing .....	27
Gambar 3. 7 Output Proses Pelabelan Data .....	29
Gambar 3. 8 Proses Filtering dan Stemming .....	30
Gambar 3. 9 Proses Algoritma Bobby Nazief dan Mirna Adriani .....	30
Gambar 3. 10 Flowchart Algoritma KNN .....	32
Gambar 3. 11 Hasil Proses Algoritma KNN .....	32
Gambar 3. 12 Diagram Alir Proses Data .....	33
Gambar 3. 13 Proses Pengujian Akurasi .....	33
Gambar 4. 1 Diagram Data .....	36
Gambar 4. 2 Grafik pengujian Variasi nilai K-NN .....	42

## DAFTAR TABEL

Tabel 3.1 Proses Sebelum Cleaning.....	22
Tabel 3.2 Proses Sesudah Cleaning.....	23
Tabel 3.3 Proses Case Folding .....	25
Tabel 3.4 Hasil Tokenizing dari Case Folding .....	27
Tabel 4.1 Confusion Matrix .....	43
Tabel 4.2 <i>Confusion Matrix</i> Algoritma KNN .....	44
Tabel 4.3 Kinerja Algoritma KNN .....	46
Tabel 4.4 Hasil lengkap perhitungan .....	47
Tabel 4.5 TPR, FPR dan Threshold .....	48



## INTISARI

Angka pengguna internet sosial media mencapai 204,7 juta jiwa pada awal tahun 2022, naik sekitar 1,03% dibandingkan awal tahun sebelumnya yang tercatat berjumlah 202,6 juta. Indonesia merupakan negara di urutan ke 5 dalam penggunaan sosial media Twitter. Banyak diskusi yang dilakukan salah satunya tentang bergabungnya dua perusahaan teknologi asal Indonesia Gojek dan Tokopedia menjadi bisnis raksasa bernama GoTo.

Metode K-Nearest Neighbors untuk klasifikasi atau klustering tweet berdasarkan tetangga terdekat pada data latih. Dengan begitu penelitian ini dapat digunakan untuk menentukan apakah tweet tentang *Initial Public Offering* (IPO) Pt GoTo Gojek Tokopedia Tbk tersebut berorientasi sentiment negative atau positif serta mengetahui presentasi akurasi dari Metode K-NN dalam klasifikasi tweet Bahasa Indonesia.

Penelitian ini menganalisis sentiment pengguna Twitter berupa data tweet dalam Bahasa Indonesia dengan metode K-NN (K-Nearest Neighbors). Pengujian menggunakan data tweet 410 tweet dengan kata kunci IPO GoTo, IPO Gojek, IPO Tokopedia, dilakukan dengan pembagian 80% data latih dan 20% data uji untuk melihat akurasi dan kinerja hasil klasifikasi. Berdasarkan pengujian diperoleh akurasi sebesar 73%.

**Kata kunci :**

**Analisis Sentimen, Algoritma KNN, K-Nearest Neighbors, Saham GoTo**

## **ABSTRACT**

*The number of social media internet users reached 204.7 million in early 2022, an increase of around 1.03% compared to the beginning of the previous year which was recorded at 202.6 million. Indonesia is the 5th country in the use of social media Twitter. Many discussions were held, one of which was about the merger of two technology companies from Indonesia, Gojek and Tokopedia, to become a giant business called GoTo.*

*The K-Nearest Neighbors method for classification or clustering tweets based on the nearest neighbors in the training data. That way this research can be used to determine whether the tweets about the Initial Public Offering (IPO) of Pt GoTo Gojek Tokopedia Tbk are oriented towards negative or positive sentiments and to know the presentation accuracy of the K-NN Method in the classification of Indonesian tweets.*

*This study analyzes Twitter user sentiment in the form of tweet data in Indonesian using the K-NN (K-Nearest Neighbors) method. Testing using tweet data of 410 tweets with the keywords GoTo IPO, Gojek IPO, Tokopedia IPO, was carried out by dividing 80% training data and 20% test data to see the accuracy and performance of the classification results. Based on the test obtained 73% accuracy.*

**Keywords :**

***Sentiment analysis, KNN Algoritm, K-Nearest Neighbors, GoTo Stock***