

**PENERAPAN DATA MINING UNTUK MEMPREDIKSI MINAT
BELI PADA MEDIA SOSIAL INSTAGRAM DENGAN
METODE K-NEAREST NEIGHBOR**

SKRIPSI



diajukan oleh
MUHAMMAD DENY RIVALDY
18.11.2431

Kepada
PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023

**PENERAPAN DATA MINING UNTUK MEMPREDIKSI MINAT
BELI PADA MEDIA SOSIAL INSTAGRAM DENGAN
METODE K-NEAREST NEIGHBOR**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



diajukan oleh

MUHAMMAD DENY RIVALDY

18.11.2431

Kepada

PROGRAM SARJANA

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2023

HALAMAN PERSETUJUAN

SKRIPSI

**PENERAPAN DATA MINING UNTUK MEMPREDIKSI MINAT
BELI PADA MEDIA SOSIAL INSTAGRAM DENGAN
METODE K-NEAREST NEIGHBOR**

yang disusun dan diajukan oleh

MUHAMMAD DENY RIVALDY

18.11.2431

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 2 Januari 2023

Dosen Pembimbing



Mardhiya Hayaty, S.Kom, M.Kom
NIK. 190302108

HALAMAN PENGESAHAN

SKRIPSI

**PENERAPAN DATA MINING UNTUK MEMPREDIKSI MINAT
BELI PADA MEDIA SOSIAL INSTAGRAM DENGAN
METODE K-NEAREST NEIGHBOR**

yang disusun dan diajukan oleh

Muhammad Deny Rivaldy

18.11.2431

Telah dipertahankan di depan Dewan Penguji
pada tanggal 17 Januari 2023

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Rumini, M.Kom
NIK. 190302246



Uyock Anggoro Saputro, M.Kom
NIK. 190302419



Mardhiyah Hayaty, S.T., M.Kom
NIK. 190302108



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 17 Januari 2023

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muhammad Deny Rivaldy
NIM : 18.11.2431

Menyatakan bahwa Skripsi dengan judul berikut:

Penerapan Data Mining Untuk Memprediksi Minat Beli Pada Media Sosial Instagram Dengan Metode K-Nearest Neighbor

Dosen Pembimbing : Mardhiya Hayaty

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 17 Januari 2023

Yang Menyatakan,



Muhamad Deny Rivaldy

HALAMAN PERSEMBAHAN

Alhamdulillah segala puji dan syukur kepada Allah SWT atas segala rahmat dan karunia telah dilimpahkan kepada penulis sehingga penulis mampu menyelesaikan skripsi dengan baik. Penulis juga mengucapkan terima kasih kepada pihak yang membantu baik secara langsung maupun tidak langsung. Skripsi ini saya persembahkan kepada:

1. Allah SWT untuk segala rahmat dan nikmat yang telah diberikan.
2. Kepada kedua orang tua yang sangat berarti bagi hidup saya. Terima kasih atas segala pengorbanan dan doa baik yang tidak pernah putus diberikan kepada saya, dan juga terima kasih telah menghadirkan saya ke dunia.
3. Keempat kakak saya (Roma, Sri, Syefira, Wahyu), adik saya (Zaky) yang selalu mendoakan dan memberikan semangat serta motivasi hingga saya dapat menyelesaikan skripsi.
4. Dosen pembimbing saya Ibu Mardhiya Hayaty, S.T., M. Kom yang telah membimbing saya dari awal sampai akhir pembuatan skripsi.
5. Dosen-dosen Universitas Amikom Yogyakarta yang telah memberikan banyak ilmu selama kuliah.
6. Kepada diri saya sendiri, terima kasih telah berjuang, terima kasih telah bekerja keras, terima kasih telah bertahan dalam segala proses meraih langkah awal untuk mewujudkan impian saya.
7. Teman-teman Kiki, Toriq, Wahyu, Ridho yang memberikan semangat untuk menyelesaikan skripsi. Alhamdulillah segala puji dan syukur kepada Allah SWT atas segala rahmat dan karunia yang telah dilimpahkan kepada penulis sehingga penulis dapat menyelesaikan tugas skripsi sebagai salah satu syarat kelulusan untuk memperoleh gelar sarjana pada Program Studi Informatika. Shalawat dan salam selalu terlimpahkan keharibaan Rasulullah Muhammad SAW.

KATA PENGANTAR

Puji Syukur penulis panjatkan kehadiran 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 dan mulia dan menuntun umatnya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Penerapan Data Mining Untuk Memprediksi Minat Beli Pada Media Sosial Instagram dengan Menggunakan Metode K-Nearest Neighbor” sebagai salah satu syarat untuk memperoleh gelar sarjana pada Program Studi Informatika.

Penulis menyadari bahwa tidaklah mudah dalam menyelesaikan skripsi ini tanpa bantuan dan bimbingan dari berbagai pihak yang senantiasa memberikan ilmu, motivasi, bantuan, bimbingan dan pengarahan dari berbagai pihak, untuk itu dengan segala hormat mengucapkan terima kasih khususnya kepada:

1. Bapak Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Mardhiya Hayaty, S. Kom, M. Kom. selaku dosen pembimbing yang telah banyak memberikan nasehat dan arahan agar penulis menyelesaikan skripsi ini dengan baik.
3. Bapak Ibu Dosen Universitas Amikom Yogyakarta yang telah memberikan ilmunya selama kuliah.
4. Keluarga dan teman-teman yang telah memberikan dukungan dan do'a kepada penulis.
5. Kepada semua pihak yang memberikan bantuan baik moril maupun materil sehingga laporan skripsi ini dapat diselesaikan.

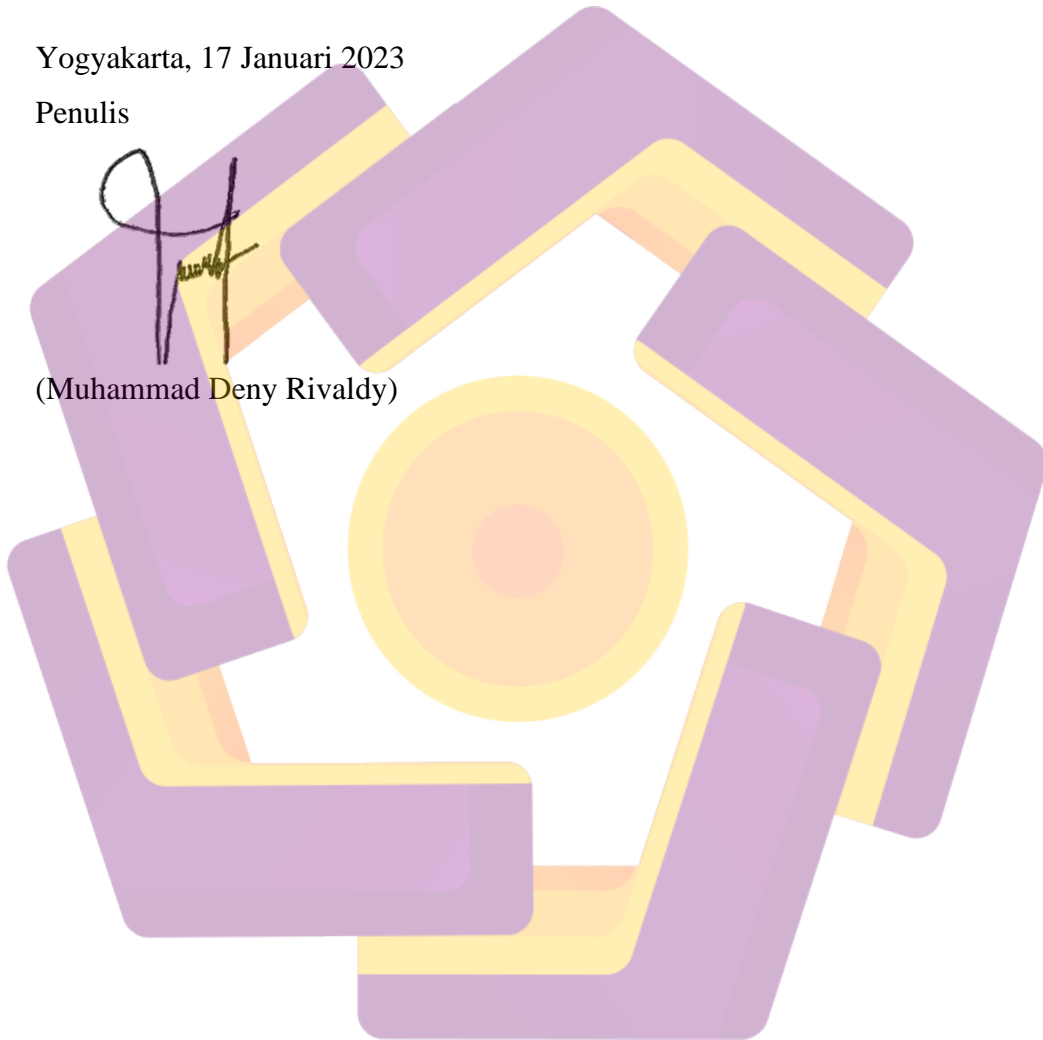
Penulis menyadari masih banyak kekurangan yang masih harus diperbaiki dari penulisan skripsi ini. Penulis menerima kritik dan saran yang membangun untuk kesempurnaan skripsi ini. semoga tulisan ini bisa bermanfaat dan mendorong kita untuk melakukan penelitian yang lebih baik.

Yogyakarta, 17 Januari 2023

Penulis



(Muhammad Deny Rivaldy)



DAFTAR ISI

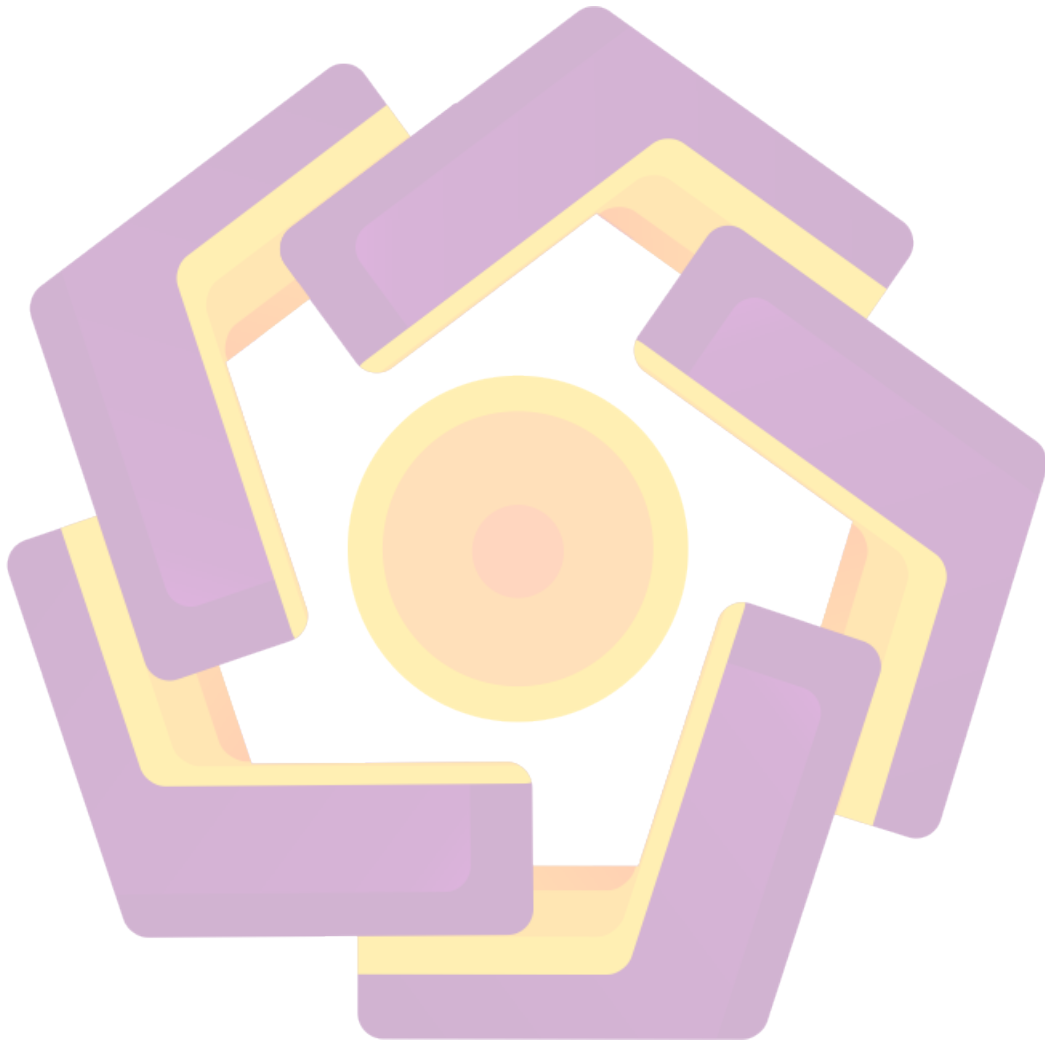
HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMBANG DAN SINGKATAN	xii
INTISARI	xiii
Abstract	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan masalah	5
1.3 Tujuan Penelitian	5
1.4 Batasan Masalah	5
1.5 Manfaat Penelitian	5
1.6 Metode Penelitian	6
1.6.1 Metode Pengumpulan Data	6
1.6.2 Metode Analisis	6
1.6.3 Metode Perancangan	6
1.6.4 Metode Implementasi	6
1.6.5 Metode Pengujian	6
BAB II TINJAUAN PUSTAKA	7
2.1 Literature Review	7
2.2 <i>Data Mining</i>	13
2.3 K -Nearest Neighbor	14
2.4 Klasifikasi	17
2.5 Preprocessing	18
2.5.1 Case Folding	18

2.5.2	Tokenizing	18
2.5.3	Stopword removal	18
2.6	Pengujian	18
2.6.1	Confusion Matrix	19
2.7	Analisis Kebutuhan	19
BAB III METODOLOGI PENELITIAN		21
3.1	Analisis Penelitian	21
3.1.1	Identifikasi Masalah	21
3.1.2	Pengumpulan Data	22
3.2	Analisis Kebutuhan	22
3.2.1	Analisis Kebutuhan Fungsional	22
3.2.2	Analisis Kebutuhan Non-Fungsional	22
3.3	<i>Preprocessing</i>	23
3.4	Pembagian Data	23
3.5	Tahap Pengujian	24
BAB IV HASIL DAN PEMBAHASAN		25
4.1	Hasil Pengumpulan Data	25
4.1.1	Implementasi Pengumpulan Data	25
4.1.2	Instagram scraper	26
4.1.3	Hasil Scraper	27
4.2	Hasil Processing Data	28
4.3	Hasil Pelabelan Data	30
4.4	Hasil Pembagian Data	31
4.5	Hasil K-Nearest Neighbor	31
4.6	Hasil Evaluasi	32
BAB V KESIMPULAN DAN SARAN		35
5.1	Kesimpulan	35
5.2	Saran	35
DAFTAR PUSTAKA		36

DAFTAR TABEL

Tabel 2. 1 Keaslian Peneliti

9



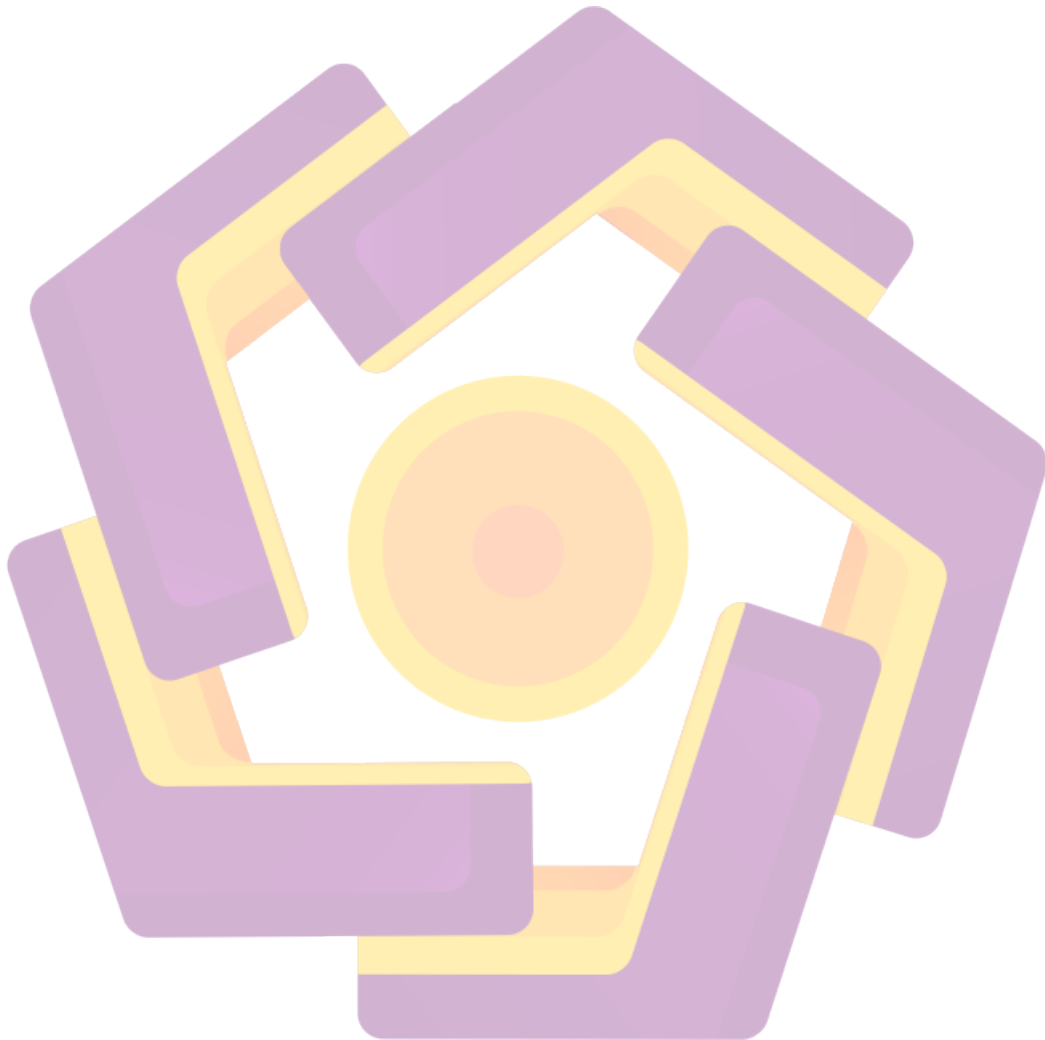
DAFTAR GAMBAR

Gambar 2. 1 Grafik penghitungan KNN.....	16
Gambar 2. 2 KNN Clasifikasi.....	18
Gambar 2. 3 Confusion Matrix.....	19
Gambar 3. 1 Alur Penelitian.....	21
Gambar 4. 1 Kode Program Authentication Access Token.....	25
Gambar 4. 2 Kickavenue.....	25
Gambar 4. 3 IG Scraper dan Email.....	26
Gambar 4. 4 IG Scraper.....	27
Gambar 4. 5 Hasil Scraper.....	27
Gambar 4. 6 Script Case Folding.....	28
Gambar 4. 7 Tahapan Upercase.....	28
Gambar 4. 8 Tahapan Lowercase.....	28
Gambar 4. 9 Script Tokenizing.....	29
Gambar 4. 10 Hasil Tokenizing.....	29
Gambar 4. 11 Import Library Stopword.....	29
Gambar 4. 12 List Stopword Remov.....	30
Gambar 4. 13 Hasil Stopword Text Final.....	30
Gambar 4. 14 Script Pelabelan Data.....	30
Gambar 4. 15 Hasil Labeling post Caption Ya dan Tidak.....	31
Gambar 4. 16 Script Pembagian Data.....	31
Gambar 4. 17 Grafik Akurasi.....	32
Gambar 4. 18 Hasil Klasifikasi K-NN tetangga terdekat.....	32
Gambar 4. 19 Akurasi Minat Ya dan Tidak.....	33
Gambar 4. 20 Confusion Matrix.....	33
Gambar 4. 21 Presisi,recall,dan F1 – Score.....	34
Gambar 4. 22 Nilai akurasi Score Caption ,Jumlah Likes.....	34

DAFTAR LAMBANG DAN SINGKATAN

KNN = K -Nearest Neighbor

IG = Instagram



INTISARI

Di era milenial saat ini, aplikasi jejaring sosial atau yang disebut platform media sosial menjadi kebutuhan tersier yang hampir dimiliki oleh setiap orang, baik itu pelajar atau pekerja, anak-anak hingga orang dewasa dan tentunya para pelaku dunia hiburan atau seni. Salah satu platform media sosial yang saat ini paling banyak diakses oleh pengguna adalah Instagram yang merupakan media sosial berbasis gambar yang menyediakan layanan berbagai foto atau video menjadi feed yang dapat diedit dengan berbagai filter dan dikelola dengan tag informasi lokasi. Pencarian platform Instagram saat ini paling banyak dicari dan diminati oleh para pengguna. Ada banyak kegunaan Instagram yang bisa dimanfaatkan pengguna, salah satunya adalah berbisnis melalui Instagram dengan menjadikan akunnya sebagai toko jualan kekinian untuk mengumpulkan produk pelanggan. Pada penelitian kali ini dilakukan dengan memprediksi minat beli pada media sosial Instagram. Sistem prediksi minat beli ini dibuat bertujuan untuk membantu pelanggan dalam memprediksi niat pembelian. Sistem ini menggunakan metode KNN atau K-Nearest Neighbor untuk mengklasifikasi objek baru berdasarkan jumlah post caption. Keakuratan terbesar adalah 99,86% untuk minat Ya atau Tidak dalam memprediksi minat beli pada media sosial Instagram.

Kata kunci: Data Mining, K-Nearest Neighbor, Klasifikasi, Instagram, Minat Beli

Abstract

In the current millennial era, social networking applications or so-called social media platforms are a tertiary need that almost everyone has, be it students or workers, children to adults and of course actors in the entertainment or arts world. One of the social media platforms that is currently the most widely accessed by users is Instagram, which is an image-based social media that provides services of various photos or videos into feeds that can be edited with various filters and managed with location information tags. The search for the Instagram platform is currently the most sought after and in demand by users. There are many uses for Instagram that users can take advantage of, one of which is doing business through Instagram by making their account a current selling shop to collect customer products. In this research, it was carried out by predicting purchase intention on Instagram social media. This buying interest prediction system was created to assist customers in predicting purchase intentions. This system uses the KNN or K-Nearest Neighbor method to classify new objects based on the number of post captions. The greatest accuracy is 99.86% for Yes or No interest in predicting buying interest on Instagram social media

Keyword: *Data Mining, K-Nearest Neighbor, Classification, Instagram, Purchase Interest*

