



PENGARUH PARAMETER K DALAM K-NEAREST NEIGHBORS TERHADAP AKURASI REKOMENDASI FILM

RAMADHAN WIRADIKUSUMA



**PROGRAM SARJANA ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**

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 IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.





PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Pengaruh Parameter K dalam *K-Nearest Neighbors* Terhadap Akurasi Rekomendasi Film” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2025

Ramadhan Wiradikusuma
G64180096

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 mengikuti kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



ABSTRAK

RAMADHAN WIRADIKUSUMA, Pengaruh Parameter K Dalam *K-Nearest Neighbors* Terhadap Akurasi Rekomendasi Film. Dibimbing oleh TOTO HARYANTO.

Dalam era digital saat ini, rekomendasi film telah menjadi salah satu fitur penting yang ditawarkan oleh berbagai platform streaming. Akurasi dalam sistem rekomendasi sangat krusial, karena dapat mempengaruhi pengalaman pengguna dan keputusan untuk menonton film tertentu. Salah satu metode yang populer dalam sistem rekomendasi adalah *K-Nearest Neighbors* (KNN). Metode ini bekerja dengan cara mencari K tetangga terdekat dari data pengguna atau item yang sedang dianalisis. Tujuan dari penelitian ini adalah untuk menganalisis pengaruh parameter K dalam penerapan algoritma KNN. Penelitian ini menggunakan dataset dari situs Kaggle yang berjudul *The Ultimate 1Million Movies Dataset*, terdiri atas 1.000.467 baris data dan 27 kolom atribut. Dataset ini mencakup berbagai informasi tentang film, di antaranya judul, genre, pemeran, rating, dan lain lain. Metode pada penelitian ini meliputi tahapan pengumpulan data, analisis data, praproses data, pelatihan, dan terakhir tahap evaluasi. Hasil dari penelitian ini adalah KNN dengan menggunakan teknik optimasi parameter *Elbow method*, *Cross-validation*, dan *GridSearchCV* untuk pencarian parameter k terbaik. Nilai akurasi optimal pada parameter k = 19, sebesar 86,71%, waktu pelatihan selama 2127,7 detik dengan menggunakan teknik optimasi *Cross-validation*.

Kata kunci: k-nearest neighbors, parameter k, sistem rekomendasi

ABSTRACT

RAMADHAN WIRADIKUSUMA. The Effect of Parameter K in K-Nearest Neighbors on Movie Recommendation Accuracy. Supervised by TOTO HARYANTO.

In today's digital era, movie recommendations have become a key feature offered. Accuracy in recommendation systems is crucial, as it directly influences user experience and viewing decisions. One popular method in recommendation systems is K-Nearest Neighbors (KNN), which operates by identifying the K closest neighbors from user or item data being analyzed. This study aims to analyze the influence of the K parameter in the implementation of the KNN algorithm. The research utilizes a dataset from Kaggle titled The Ultimate 1Million Movies Dataset consisting of 1,000,467 entries and 27 attribute columns. The dataset includes various movie details such as title, genre, cast, rating, and more. The methodology involves data collection, data analysis, preprocessing, training, and evaluation. The study employs parameter optimization techniques, including the Elbow method, Cross-validation, and GridSearchCV, to determine the optimal K value. The results indicate that the optimal accuracy 86,71%, total training time for this configuration was 2127,7 seconds is achieved at K = 19 using the Cross-validation optimization.

Keywords: k-nearest neighbors, parameter k, recommendation system



©Hak cipta milik IPB University

IPB University

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 IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2025
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.



**PENGARUH PARAMETER K DALAM K-NEAREST
NEIGHBORS TERHADAP AKURASI
REKOMENDASI FILM**

RAMADHAN WIRADIKUSUMA

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Ilmu Komputer

**PROGRAM SARJANA ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**

Tim Pengaji pada Ujian Skripsi:

- 1 Firman Ardiansyah, S.Kom., M.Si.
- 2 Muhammad Asyhar Agmalaro S.Si., M.Kom.

©Hak cipta milik IPB University

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 IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Judul Skripsi : Pengaruh Parameter K dalam *K-Nearest Neighbors* Terhadap Akurasi Rekomendasi Film

Nama : Ramadhan Wiradikusuma
NIM : G64180096

Disetujui oleh

Pembimbing:

Dr. Toto Haryanto, S.Kom., M.Si.

Diketahui oleh

Ketua Program Sarjana Ilmu Komputer:

Dr. Sony Hartono Wijaya, S.Kom., M.Kom.
NIP. 19810809 200812 1 002

Tanggal Ujian:
22 Juli 2025

Tanggal Lulus:



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah *subhanaahu wa ta'ala* atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Judul yang dipilih dalam penelitian yang dilaksanakan sejak bulan Januari 2025 sampai bulan Juli 2025 ini ialah “Pengaruh Parameter K Dalam *K-Nearest Neighbors* Terhadap Akurasi Rekomendasi Film”.

Terima kasih penulis ucapan kepada dosen pembimbing, Bapak Dr. Toto Haryanto, S.Kom., M.Si. yang telah membimbing dan banyak memberi saran. Ucapan terima kasih juga disampaikan kepada dosen penguji, Bapak Firman Ardiansyah, S.Kom., M.Si. dan Bapak Muhammad Asyhar Agmalaro S.Si., M.Kom. yang telah memberikan saran dan masukan untuk kesempurnaan skripsi, pembimbing akademik, moderator seminar, dan teman-teman kampus IPB University yang telah mewarnai masa-masa kuliah penulis di kampus terbaik bangsa. Ungkapan terima kasih dan doa juga disampaikan kepada almarhum/ah kedua orang tua tercinta ayah Usma Sumiadi dan mamah Minarni yang selalu memberikan dukungan, doa, nasihat, dan kasih sayangnya. Terima kasih kepada saudari kandung kakak Evelina dan Nusantari yang selalu memberikan motivasi dan semangat. Terima kasih kepada Salwa yang selalu setia berada di samping penulis. Terima kasih kepada keluarga besar saudara/i yang selalu mendukung kesuksesan penulis. Terima kasih kepada teman-teman rumah, vespa, dan semasa sekolah yang selalu menemani kehidupan penulis. Terima kasih kepada rekan-rekan kerja Direktorat PKPS yang selalu mendukung untuk kelancaran kuliah penulis.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2025

Ramadhan Wiradikusuma



DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

PENDAHULUAN

- 1.1 Latar Belakang
- 1.2 Rumusan Masalah
- 1.3 Tujuan
- 1.4 Manfaat
- 1.5 Ruang Lingkup

TINJAUAN PUSTAKA

- 2.1 Sistem Rekomendasi
- 2.2 Algoritma *K-Nearest Neighbors*
- 2.3 *Elbow Method*
- 2.4 *Cross-validation*
- 2.5 *GridSearchCV*

METODE

- 3.1 Data Penelitian
- 3.2 Tahapan Penelitian
- 3.3 Pengumpulan Data
- 3.4 Eksplorasi Data
- 3.5 Praproses Data
 - 3.5.1 *Encoding*
 - 3.5.2 Pembersihan Data dan Seleksi Fitur
 - 3.5.3 Pembagian Data
 - 3.5.4 Standardisasi
- 3.6 Pelatihan KNN
- 3.7 Evaluasi KNN
- 3.8 Peralatan Penelitian

HASIL DAN PEMBAHASAN

- 4.1 Praproses Data
 - 4.1.1 *Encoding*
 - 4.1.2 Pembersihan Data dan Seleksi Fitur
 - 4.1.3 Pembagian Data
 - 4.1.4 Standardisasi
- 4.2 Pelatihan KNN
- 4.3 Evaluasi KNN
- 4.4 Perbandingan Penelitian

SIMPULAN DAN SARAN

- 5.1 Simpulan
- 5.2 Saran

x

x

x

1

1

2

2

2

2

3

3

3

5

6

7

8

8

8

8

8

8

9

10

10

10

10

10

10

11

12

13

13

13

13

14

15

15

16

17

19

19

19

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 mengugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



DAFTAR PUSTAKA	20
LAMPIRAN	24
RIWAYAT HIDUP	27

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 IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



1	Deskripsi detail fitur dataset	9
2	Hasil perhitungan korelasi <i>pearson</i>	14
3	Penyebaran data statistik	14
4	Distribusi data untuk tiap kelas	15
5	Percobaan ke 1 sampai 5	15
6	Metrik evaluasi hasil percobaan	17
7	Perbandingan penelitian KNN terdahulu	17

DAFTAR GAMBAR

1	Ilustrasi logika proses sistem rekomendasi (Ferreira <i>et al.</i> 2020)	3
2	Ilustrasi Perhitungan jarak <i>euclidean</i> antara 2 titik (Bansal <i>et al.</i> 2022)	4
3	Ilustrasi algoritma KNN untuk klasifikasi (Bansal <i>et al.</i> 2022)	5
4	Ilustrasi hasil perhitungan <i>Elbow method</i> (Dewi dan Pramita 2019)	6
5	Ilustrasi konsep <i>Cross-validation</i> pada penelitian (Nti <i>et al.</i> 2021)	6
6	Ilustrasi metrik <i>GridsearchCV</i> pada penelitian (Sukamto <i>et al.</i> 2023).	7
7	Tahapan penelitian	8
8	(a) Grafik optimasi parameter menggunakan <i>Elbow method</i> , (b) grafik <i>Cross-validation</i> , dan (c) grafik <i>GridSearchCV</i>	16

DAFTAR LAMPIRAN

1	Hasil perbandingan dataset 1(a) sebelum dan 1(b) sesudah dilakukan praproses <i>encoding</i>	25
2	Matriks korelasi fitur dengan target	26