

Skripsi Fisika

**IMPLEMENTASI ALGORITMA *K-NEAREST NEIGHBORS* UNTUK
MENGKLASIFIKASI BINTANG VARIABEL**

**ANDI YUSRIANDI PRATAMA
H211 15 304**



**DEPARTEMEN FISIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2022**

**IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR UNTUK
MENGKLASIFIKASI BINTANG VARIABEL**

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
Pada Program Studi Fisika Departemen Fisika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Hasanuddin**



ANDI YUSRIANDI PRATAMA

H21115304

**DEPARTEMEN FISIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR**

2022

LEMBARAN PENGESAHAN

**IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBORS UNTUK MENGLASIFIKASI
BINTANG VARIABEL**

Disusun dan Diajukan oleh :

ANDI YUSRIANDI PRATAMA

H21115304



Telah dipertahankan di hadapan panitia ujian yang dalam rangka Penyelesaian Studi Sarjana Program Studi Fisika Fakultas Matematika dan Ilmu Pengertahuan Alam Universitas Hasanuddin pada tanggal 16 Desember 2022 dan dinyatakan telah memenuhi syarat kelulusan

Pembimbing Utama

Eko Juarlin, S.Si, M.Si

NIP. 19811106 200812 1 002

Pembimbing Pertama

Nur Hasanah, S. Si, M. Si

NIP. 19831122 200912 2001

Ketua Departemen

Prof. Dr. Arifin, M.T

NIP. 19670520 199403 1 002

PERNYATAAN KEASLIAN

Dengan ini saya menyatakan bahwa skripsi ini merupakan karya orisinil saya dan sepanjang sepengetahuan saya tidak memuat bahwa yang pernah dipublikasi atau telah ditulis oleh yang lain dalam rangka tugas akhir suatu gelar akademis di Universitas Hasanuddin atau telah dikutip sesuai kaidah ilmiah yang berlaku. Saya juga menyatakan bahwa skripsi ini merupakan hasil karya saya sendiri dan dalam batas dibantu oleh pihak pembimbing.

Penulis

Andi Yusriandi Pratama

ABSTRAK

Penelitian ini menggunakan data Bintang Variabel yang berjumlah 187.662 data yang diambil dari *ASAS-SN Variabel Star Database* yang diklasifikasi menggunakan *algoritma machine learning*. Dilakukan normalisasi data *Z-score* dan proses *pre-processing*. Terdapat 9 atribut yaitu *raj2000*, *dej2000*, *mean_vmag*, *amplitude*, *period*, *phot_g_mean_mag*, *phot_pb_mean_mag*, *phot_pb_mean_mag*, *phot_rb_mean_mag*, dan *parallax*. Dan 4 *Variable type* yang digunakan terdiri dari *Cataclysmic Variable*, *Eclipsing Variable*, *Pulsating Variable*, dan *Rotational Variable*. Algoritma *K-Nearest Neighbor* merupakan algoritma *supervised learning* pada *machine learning* adalah sebuah metode untuk mengklasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Jarak pada objek dihitung dengan metode perhitungan jarak *manhattan*, *euclidean*, dan *minkowski*. Penelitian ini mencakup pengukuran performa (akurasi, presisi, recall, dan f-measure). Klasifikasi menghasilkan nilai akurasi rata dari masing-masing perhitungan *manhattan* 82.65%, *euclidean* 82.66% dan *minkowski* 82.66%.

Kata Kunci: Bintang Variabel, *K-Nearest Neighbor*, *Z-Score*, ASAS-SN

Abstract

The 187,662 variable stars used in this study were categorised using a machine learning algorithm using data from the ASAS-SN Variable Star Database. Zscore normalization and pre-processing were done. There are 9 attributes, namely raj2000, dej2000, mean_vmag, amplitude, period, phot_g_mean_mag, phot_pb_mean_mag, phot_pb_mean_mag, phot_rb_mean_mag, and parallax. Cataclysmic Variable, Eclipsing Variable, Pulsating Variable, and Rotational Variable are the four types of variables that are employed. A approach for classifying things based on learning data that is closest to the item is the K-Nearest Neighbor algorithm, which is a type of supervised learning in machine learning. euclidean, minkowski, and manhattan methods for calculating distance. There are performance metrics in this study (accuracy, precision, recall, and f-measure). Manhattan calculations produce an average accuracy value of 82.65%, Euclidean calculations produce an average accuracy value of 82.66%,and 82.66% for Minkowski calculations.

Keywords: *Variable Star, K-Nearest Neighbor, Z-Score, ASAS-SN.*

KATA PENGANTAR

Puji syukur kepada Allah SWT Tuhan Yang Maha Esa, Atas limpahan karunia, rahmat, nikmat serta hidayah yang tak pernah berhenti mengalir dari-Nya sehinggalah penulis dapat menyelesaikan skripsi yang berjudul “**Implementasi Algoritma K-Nearest Neighbor untuk mengklasifikasi Bintang Variabel**” yang merupakan tugas akhir untuk melengkapi persyaratan mencapai gelar sarjana Sains Departemen Fisika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin. Shalawat dan salam tak lupa penulis hanturkan kepada **Nabi Muhammad SAW**. Nabi yang menjadi teladan bagi umat dalam berakhlak, berusaha, dan berdoa.

Penulis menyadari dalam penulisan skripsi ini tidak akan dapat saya selesaikan tanpa doa dan dukungan dari berbagai pihak baik secara moril dan materil. Oleh karena itu, penulis menyampaikan rasa terima kasih yang setulus-tulusnya serta penghargaan yang setinggi-tingginya untuk orang tua penulis, Ayahanda dan Ibunda tercinta **Andi Agussalim** dan **Yusba**, yang telah membesarkan dan mendidik dengan penuh cinta, kesabaran dan kasih sayang, serta dukungan dan doa yang tulus tanpa henti yang takkan pernah bias penulis balas.

Penghargaan yang tulus dan ucapan terima kasih dengan penuh keikhlasan juga penulis ucapkan kepada:

1. **Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.**, selaku Rektor Universitas Hasanuddin beserta seluruh jajarannya.
2. **Bapak Dr. Eng. Amiruddin, M.Si.** selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. **Bapak Prof. Dr. Arifin, M.T** selaku Ketua Departemen Fisika, segenap Dosen Pengajar dan Staf Departemen Fisika yang telah membekali ilmu dan kemudahan-kemudahan kepada Penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.
4. **Almarhumah Ibu Nur Hasanah, S.Si, M.Si** dan **Bapak Eko Juarlin, S.Si, M.Si.** selaku Pembimbing Utama yang telah ikhlas meluangkan waktu dan

pemikirannya untuk memberikan arahan, pengetahuan, motivasi dan bimbingan telah diberikan kepada Penulis.

5. **Bapak Prof. Dr. rer-nat Wira Bahari Nurdin** dan **Bapak Heryanto, S.Si, M.Si.** selaku Tim Penguji atas saran, kritikan, dan masukan yang membangun dalam penyempurnaan penyusunan skripsi ini serta waktu yang telah diberikan kepada Penulis.
6. Keluarga besar **MIPA 2015** dan **F15IKA**, terima kasih terus menemani masa mahasiswa, memberikan pelajaran, motivasi, canda, dan tawa yang menjadi pengalaman hidup bagi Penulis
7. **Keluarga Mahasiswa FMIPA Unhas** terkhusus anggota keluarga **Himafi FMIPA Unhas**, terima kasih atas ilmu, pembelajaran hidup, dan telah menjadi keluarga selama penulis kuliah di Universitas Hasanuddin.
8. **Keluarga Besar KPA Omega Himafi Unhas**, terkhusus kepada teman-teman **DIKSAR XXII**.
9. Terkhusus, Partner saya sejak 2013 **Jihan Amanda** sebagai *support system* Penulis.
10. Semua pihak yang telah banyak berpartisipasi, baik secara langsung maupun tidak langsung yang tak sempat penulis sebutkan satu per satu. Terima kasih untuk segala bantuan dan dukungannya.

Dalam penulisan dan penyusunan Skripsi ini, Penulis menyadari bahwa masih jauh dari kata sempurna. Oleh karena itu, dengan segala kerendahan hati penulis memohon maaf atas kesalahan yang dilakukan. Semoga skripsi ini dapat memberikan manfaat bagi pembaca dan semua pihak khususnya teman-teman yang berhubungan dengan Fisika.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Makassar 16 Desember 2022

Andi Yusriandi Pratama

DAFTAR ISI

SAMPUL1

HALAMAN SAMPUL2

LEMBARA PENGESAHAN Error! Bookmark not defined.

PERNYATAAN KEASLIAN Error! Bookmark not defined.

ABSTRAK5

Abstract6

KATA PENGANTAR7

DAFTAR ISI9

DAFTAR GAMBAR11

DAFTAR TABEL12

DAFTAR LAMPIRAN13

BAB I PENDAHULUAN14

I.1 Latar Belakang14

I.2 Rumusan Masalah15

I.3 Tujuan15

BAB II TINJAUAN PUSTAKA16

II.1 Definisi Bintang16

II.2 Bintang Variabel (*Variable Stars*)16

II.2.1 Variable Intrinsik17

II.2.2 Variable Ekstrinsik19

II.3 *K-Nearest Neighbor*21

BAB III METODOLOGI PENELITIAN24

III.1 Studi literatur24

III. 2 Sumber Data24

III.3 *Pre-Processing/Cleaning* (Pembersihan Data)24

III.4 Transformasi Data25

III.5 *Split* dan Visualisasi Data26

III.5 Normalisasi data26

III.6 Implementasi Metode27

III.7 Evaluasi Performa27

III.8.1 Bagan Alir Metode *K-Nearst Neighbor*29

III.8.2 Bagan Alir Penelitian	30
<i>BAB IV HASIL DAN PEMBAHASAN</i>	31
IV.1 Visualisasi Data	31
IV.2 Klasifikasi Bintang Variabel	32
IV.2.1 Normalisasi Data dengan Z-score	32
IV.2.2 Hasil K-Nearest Neighbors	34
IV.2 Evaluasi Hasil	35
<i>BAB V KESIMPULAN DAN SARAN</i>	41
<i>DAFTAR PUSTAKA</i>	42
<i>LAMPIRAN</i>	46

DAFTAR GAMBAR

Gambar 2.1 Kurva Cahaya Bintang Variabel.....	3
Gambar 2.2 Klasifikasi Bintang Variabel.....	4
Gambar 2.3 (a) Chepeid (b) Kurva Cahaya Chepeid.....	5
Gambar 2.4 Ilustrasi sebuah system cataclysmic variable.....	6
Gambar 2.5 Ilustrasi dan kurva cahaya bintang gerhana ganda.....	7
Gambar 2.6 Metode K-Nearest Neighbor.....	9
Gambar 3.1 Diagram Alir metode K-Nearest Neighbor.....	15
Gambar 3.2 Diagram Alir Penelitian.....	17
Gambar 4.1 Persentasi Bintang Variabel dalam dataset.....	18
Gambar 4.2 Split dataset (a) Data Training (b) Data Testing.....	19
Gambar 4.3 Grafik perbandingan akurasi.....	27

DAFTAR TABEL

Tabel 3.1 Atribut dalam database ASAS-SN.....	12
Tabel 3.2 Transformasi data objek.....	13
Tabel 3.3 Table Confusion Matrix 2 kelas.....	15
Tabel 4.1 Contoh data yang dinormalisasi.....	19
Tabel 4.2 Nilai rata-rata dan Standar Deviasi dari atribut.....	20
Tabel 4.3 Hasil Normalisasi <i>Z-Score</i>	20
Tabel 4.4 Data Latih Sampel.....	21
Tabel 4.5 Data Uji Sampel.....	21
Tabel 4.6 Perhitungan jarak pada data latih ke data uji.....	22
Tabel 4.7 Performa model pada perhitungan jarak manhattan.....	23
Tabel 4.8 Performa model pada perhitungan jarak Euclidean.....	24
Tabel 4.9 Performa model pada perhitungan jarak minkowski.....	25
Tabel 4.10 Hasil akurasi.....	26

DAFTAR LAMPIRAN

Lampiran 1 Sampel data bintang variable dengan normalisasi Z-score.....	33
Lampiran 2 Syntax Simulasi dengan <i>library Python</i>	35

BAB I

PENDAHULUAN

I.1 Latar Belakang

Bintang Variabel adalah bintang yang kecerahannya bervariasi [1]. Bintang variable digunakan untuk mempelajari populasi dan struktur bintang. Bintang Variabel terbagi menjadi dua kategori yaitu bintang variabel intrinsik dan bintang variable ekstrinsik. Bintang variable intrinsik adalah bintang yang variabilitasnya disebabkan oleh perubahan sifat fisik bintang itu sendiri, bintang variable kategori ini dibagi menjadi 2 subkelompok yaitu bintang berdenyut (*pulsating variable*) [1] dan bintang kataklismik (*cataclysmic variable*) sedangkan bintang variable ekstrinsik adalah bintang di mana variabilitasnya disebabkan oleh sifat eksternal seperti rotasi atau gerhana, kategori ini juga terbagi menjadi dua subkelompok yaitu *eclipsing binaries* dan *rotational variable* [3].

Data astronomi setiap tahunnya meningkat secara eksponensial sehingga mustahil bagi para astronomi atau para peneliti melakukan verifikasi dan memberi label pada setiap objek individu secara manual. Pelabelan ini dapat memakan waktu yang lama dan kemungkinan terjadi bias sangat besar. Oleh karena itu diperlukan metode dalam analisis dan eksplorasi data yang efektif dan akurat. Solusi untuk masalah ini yang paling mungkin adalah melibatkan algoritma *machine learning* [4].

Penerapan machine learning sudah dilakukan sebelumnya, seperti klasifikasi galaksi, kuasar dan bintang [5], morfologi galaksi [6], verifikasi gelombang gravitasi [7] dan masih banyak lagi. Termasuk dalam penyusunan katalog ASASSN bintang variabel [8] yang memanfaatkan machine learning dalam proses klasifikasi bintang. Lianping Tu, et al membandingkan algoritma *machine learning K-Nearest Neighbors* (K-NN), dan *Support Vektor Machine* (SVM) pada penelitiannya yang mengklasifikasikan galaxy dan kuasar, dalam perbandingan tersebut didapat rata-rata pengklasifikasian galaxy dan kuasar pada metode K-NN sebesar 97,94 % dan pada SVM sebesar 96.46 % [9]. Penelitian sebelumnya yang terkait juga menerapkan beberapa algoritma *K-Nearest*

Neighbor algorithm (K-NN), Decision Tree (DT), Random forest (RF) and Support Vektor Machine (SVM) dengan rata-rata akurasi 98% [10],

Berdasarkan hal yang dijelaskan diatas, maka penulis akan melakukan penelitian mengenai klasifikasi bintang variable menggunakan algoritma *machine learning K-Nearest Neighbor*. Beberapa penelitian telah menunjukkan kelebihan dalam menggunakan algoritma K-NN dibandingkan algoritma machine learning SVM [9]. Penelitian ini diharapkan dapat membuat model machine learning dalam mengklasifikasikan keempat subkelompok dari bintang variable yaitu *pulsating variable*, *cataclysmic variable*, *eclipsing binaries* dan *rotational variable*.

I.2 Rumusan Masalah

Rumusan masalah dari penelitian ini antara lain

1. Bagaimana mengolah data ASAS-SN dalam klasifikasi bintang variable dengan algoritma *K-Nearest Neighbor*?
2. Bagaiman tingkat akurasi dari penerapan algoritma *K-Nearest Neighbor*?

I.3 Tujuan

Tujuan dilaksanakannya penelitian ini antara lain

1. Menghasilkan algoritma machine learning dengan bahasa programan phyton yang mampu mengklasifikasikan bintang variable menggunakan *K-Nearest Neighbor*.
2. Menganalisis tingkat akurasi dari algoritma *K-Nearest Neighbor*.

BAB II

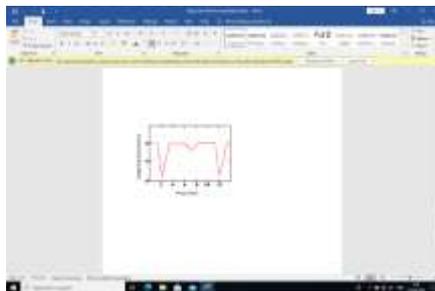
TINJAUAN PUSTAKA

II.1 Definisi Bintang

Bintang merupakan hasil dari kondensasi dari awan nebula. Akibat dari peristiwa kondensasi tekanan dari dalam awan melakukan ekspansi dengan nilai tekanan yang lebih kecil dari gaya gravitasi yang melingkupi proses kondensasi. Kemudian awan nebula yang pada awalnya menyatu dalam satu gumpalan besar terpecah dan mengalami fragmentasi [11]. Menurut ilmu astronomi, bintang adalah semua benda masif (bermassa antara 0,08 hingga 200 massa matahari) yang sedang dan pernah melangsungkan pembangkitan energi melalui reaksi fusi nuklir. Komposisi bintang terutama terdiri dari 71% hidrogen, 27% helium, dan 2% unsur berat lainnya [12]. Bintang-bintang yang terlihat di langit malam semuanya terletak dalam Galaksi Bima Sakti [13].

II.2 Bintang Variabel (*Variable Stars*)

Ketika melihat ke langit malam, sebagian besar beranggapan bahwa bintang tidak berubah. Selain berkelap-kelip karena efek atmosfer, bintang tampak tetap dan konstan bagi mata. Namun, bintang mengalami perubahan kecerahan dari waktu ke waktu. Beberapa menunjukkan perilaku periodik, cerah dengan cepat kemudian berkurang kecerahannya secara perlahan hanya untuk terulang kembali. Proses ini merupakan contoh dari bintang variabel. Bintang variabel adalah bintang dengan kecerahan yang bervariasi [1].



Gambar 2.1 Kurva Cahaya Bintang Variabel [14].

Kurva diatas merupakan kurva salah satu bintang variabel yaitu *eclipsing binaries* atau biasa disebut bintang gerhana ganda pada kurva cahaya diatas didapat bahwa terjadi perubahan cahaya yang signifikan pada kurung waktu 10 hari, Para astronom menyebut bahwa sistem biner memiliki periode waktu orbit selama 10 hari [14]. Perubahan yang terjadi pada bintang gerhana ganda juga terjadi pada bintang variabel lainnya tetapi dengan variabilitas yang berbeda, maka dari itu para peneliti mengelompokkan bintang variabel berdasarkan variabilitasnya. Dapat dilihat pada Gambar 2.2 yang menunjukkan pengelompokkan dari klasifikasi bintang variabel.



Gambar 2.2 Klasifikasi Bintang Variabel

II.2.1 Variable Intrinsik

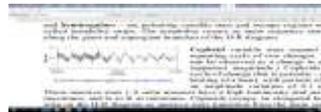
Variabel intrinsik adalah variabel di mana variabilitasnya disebabkan oleh perubahan sifat fisik bintang itu sendiri. Kategori ini dapat dibagi menjadi dua subkelompok:

- a) Pulsating variables, bintang berdenyut adalah bintang yang secara fisik mengembang dan mengerut [1]. Proses pulsating pada bintang disebabkan oleh ekspansi dan kontraksi periodik dari lapisan permukaan bintang [15]. Dalam [1], dijelaskan bahwa di bagian dalam bintang terdapat medan gravitasi sebagai akibat dari adanya massa bintang yang menyatukan semua komponen penyusun bintang. Terdapat pula tekanan yang membuat bintang mengembang. Pada bintang non-variabel, kedua kuantitas ini seimbang sehingga bintang akan tetap pada ukurannya. Pada bintang variabel yang

telah mengalami kontraksi melalui beberapa gangguan, saat suhunya naik maka atom helium di lapisan luar bintang menjadi terionisasi. Helium yang terionisasi menghentikan panas dan cahaya dari dalam bintang agar tidak keluar semudah sebelumnya. Artinya, suhu dan tekanan meningkat lebih cepat di dalam bintang. Ketika tekanan menjadi lebih besar daripada gaya gravitasi, maka bintang akan mengembang. Lalu jika bintang mencapai keadaan setimbang, tekanan akan menurun dan gas mendingin serta ion helium akan bergabung kembali menjadi atom utuh. Dalam keadaan ini, bintang akan jauh lebih transparan dan mendingin lebih cepat. Hal ini menandakan tekanan gas menurun secara tajam sehingga gaya gravitasi akan jauh lebih kuat dan bintang kembali mengerut. Gambar 2.3 merupakan salah satu objek dan kurva cahaya dari bintang berdenyut.



(a)



(b)

Gambar 2.3 (a) Cepheid [16] kurva cahaya Cepheid [15]

- b) Bintang variable kataklismik atau *cataclysmic variable* (CV) adalah suatu sistem bintang ganda yang terdiri dari komponen primer bintang katai putih dan pasangannya adalah sebuah bintang deret utama. Kedua komponen dalam sistem CV hanya berjarak beberapa radius matahari. Bintang sekunder pada sistem CV merupakan bintang yang telah berevolusi memenuhi roche lobe-nya dan memberikan transfer massa ke bintang primer. Ada beberapa jenis sistem CV dan masing-masing memiliki keunikan tersendiri. Dwarf novae, recurrent novae dan classical novae merupakan tipe sistem CV yang mengalami perubahan kecerlangan akibat outburst. Outburst ini berasal dari aliran massa yang tidak stabil pada piringan dan menyebabkan reaksi

termonuklir di permukaan bintang katai putih. Beberapa sistem lain dikategorikan sebagai polar CV karena bintang katai putihnya memiliki medan magnet yang kuat [3].



Gambar 2.4 Ilustrasi sebuah sistem *cataclysmic variable* [3].

Terjadi transfer massa dari bintang sekunder yang telah memenuhi *roche lobe* ke bintang primer (STScI). Materi yang ditransfer membawa momentum sudut dan bergerak secara spiral menuju bintang katai, membentuk sebuah cincin di sekeliling bintang katai putih. Karena ada banyak materi yang saling bergesekan maka terjadi transfer momentum sudut dari bagian dalam cincin ke arah luar, cincin akan menyebar dan membentuk piringan yang dikenal sebagai accretion disk. Selama bergerak spiral materi melepaskan energi gravitasional yang dapat memanaskan piringan hingga 3.000 – 100.000 K.

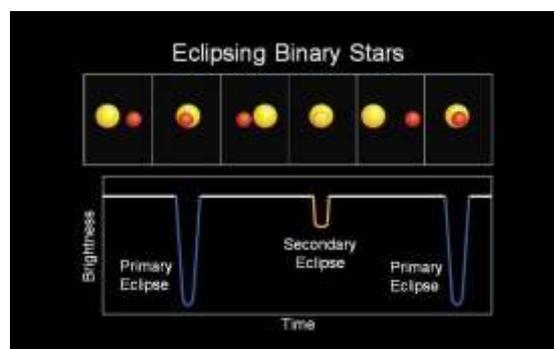
II.2.2 Variable Ekstrinsik

Variable ekstrinsik adalah bintang di mana variabilitasnya disebabkan oleh sifat eksternal seperti rotasi atau gerhana. Ada dua subkelompok utama:

- a) *Eclipsing binaries*, bintang yang jari-jarinya secara bergantian mengembang dan mengerut sebagai bagian dari proses penuaan evolusioner alaminya. Gerhana bintang biner (EBs) berfungsi sebagai landasan astrofisika bintang. Peran unik mereka yang penting dalam menentukan parameter bintang fundamental dan jarak serta memberikan tes ketat untuk model stellarevolution [17] telah dihargai secara luas. Sementara prinsip-prinsip dasar yang mengatur pemodelan EB sederhana (mekanika Newton dan pertimbangan geometri lurus ke depan), sejumlah besar komplikasi yang timbul dari efek yang lebih halus: distorsi permukaan, variasi intensitas karena penggelapan gravitasi, penggelapan tungkai, refleksi, dan penonjolan

permukaan. Seiring dengan meningkatnya presisi data, begitu pula jumlah seluk-beluk yang perlu diperhitungkan oleh model yang andal.

Kontribusi bintang biner dan, khususnya, gerhana biner (EBs) untuk astrofisika tidak dapat dilebih-lebihkan. EB bisa memberikan pengukuran massa dan jari-jari dasar untuk bintang komponen. Pengukuran massa dan radius ini pada gilirannya memungkinkan untuk tes akurat model evolusi bintang [18]. Dalam kasus di mana pengukuran kecepatan radial (RV) berkualitas tinggi ada untuk kedua bintang di EB, luminositas dihitung dari jari-jari absolut dan suhu efektif dapat menyebabkan penentuan jarak. Memang, EB menjadi banyak digunakan untuk menentukan jarak ke Awan Magellan, M31, dan M33 [18]. Sampel besar berguna untuk menentukan sifat statistik dan untuk menemukan binari langka yang mungkin menyimpan fisik signifikansi (misalnya, binari dengan bintang bermassa sangat rendah, binari dengan bintang dalam tahap evolusi berumur pendek, sangat biner eksentrik yang menunjukkan gerakan apsidal besar, dll.). Katalog EB dari survei berbasis darat mengalami berbagai bias pengamatan seperti akurasi terbatas per pengukuran individu, fungsi jendela yang kompleks (misalnya, pengamatan dari survei berbasis darat hanya dapat dilakukan pada malam hari dengan langit cerah dan pada musim tertentu). Pada Gambar 2.5 berikut mengilustrasikan proses terjadinya perubahan kecerahan pada bintang berhana ganda.



Gambar 2.5 Ilustrasi dan kurva cahaya bintang gerhana ganda [19]

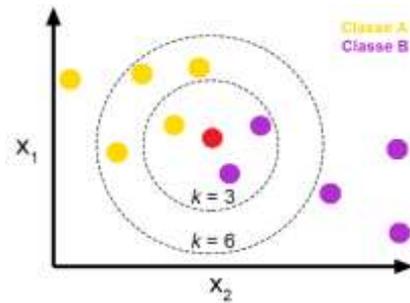
- b) *Rotating variables*, bintang yang variabilitasnya disebabkan oleh fenomena yang berkaitan dengan rotasinya. Contohnya adalah bintang dengan "bintik

matahari" ekstrem yang memengaruhi kecerahan semu atau bintang yang memiliki kecepatan rotasi cepat yang menyebabkannya berbentuk elips. Karakteristik tipe ROT adalah variasi amplitudo pada skala waktu bulan. Sering ada efek pemukulan dan travelling fitur dalam kurva cahaya, yang merupakan fitur karakteristik variabel tipe akhir dengan bintik bintang. Kami menemukan 45 bintang yang sepertinya variabel rotasi murni atau rotasi dengan jenis variasi lainnya. Kami menganggap kehadiran harmonik sebagai indikasi utama ROT Kelas [20].

II.3 *K-Nearest Neighbor*

K-Nearest Neighbor (K-NN) adalah salah satu dari banyaknya algoritma yang paling banyak digunakan dalam data mining pada proses klasifikasi [21]. K-NN memiliki daya tarik tersendiri yang membuat algoritma ini sering digunakan. Hal ini disebabkan, K-NN merupakan teknik klasifikasi yang sederhana, intuitif, dan mudah disesuaikan disegala domain permasalahan. Selain itu K-NN juga memiliki kerja yang cukup baik.

K-NN adalah algoritma supervised yang melakukan klasifikasi pada objek yang belum diketahui kelasnya dan mencari objek terdekat yang memiliki kesamaan kelas. Pada proses penentuan label atau kelas bagi data yang baru, K-NN menggunakan sistem vote majority yang sederhana [22]. Dimana kelas mayoritas dari K tetangga yang paling dekat akan dijadikan sebagai kelas terhadap data uji yang baru. K-NN bekerja dengan dengan prinsip melihat tetangga terdekat [23]. Nilai K yang terbaik untuk K-NN tergantung pada data. Secara umum, nilai K yang tinggi akan mengurangi efek noise pada klasifikasi, tetapi membuat batasan antara setiap klasifikasi menjadi lebih kabur. Nilai K yang bagus dapat dipilih dengan optimasi parameter, jadi untuk mendapatkan nilai K yang optimal biasanya dilakukan pengujian secara berulang dengan nilai K yang berbeda [24].



Gambar 2.6. Metode K-Nearst Neighbors

Pada gambar diatas dijelaskan jika terdapat dua kelas atau label dan menggunakan nilai $k=3$ dan $k=6$. Pada $k=3$ terdapat 3 objek terdekat dengan objek yang belum diketahui (merah) kelasnya. 2 objek ungu dan 1 objek kuning. Maka, objek yang tidak memiliki kelas akan memiliki kelas ungu karena objek ungu adalah objek terbanyak dan terdekat dengan nilai $k=3$, begitu pula pada $k=6$ terdapat 4 kelas kuning dan 2 kelas ungu makan objek akan memiliki kelas kuning berdasarkan tetangga terdekatnya.Min

Menurut Prasath *et. al* [25] performa K-NN tergantung dari teknik perhitungan jarak yang digunakan dengan mempertimbangkan Nilai K. Terdapat 3 cara paling populer dalam menghitung jarak tetangga terdekat pada metode K-NN sebagai berikut

1. Jarak *Euclidean* dinyatakan pada persamaan 2.1, yang mendefinisikan total perbedaan kuadrat pada masing-masing atribut dua data.

$$D(d_x, d_y) = \sqrt{(f | |x,1 - f_{y,1})^2 + \dots + (f | |x,j - f_{y,j})^2} \quad (2.1)$$

2. Jarak *Manhattan*, perhitungan yang menggunakan selisih absolut pada masing-masing atribut.

$$D(d_x, d_y) = \max(|f_{x,1} - f_{y,1}| + \dots + |f_{x,j} - f_{y,j}|) \quad (2.2)$$

3. Jarak *Minkowski*, mempunyai bentuk perhitungan yang hampir mirip seperti *Euclidean* dan *Manhattan*, namun pangkat p , dengan $1 < p < 2$.

$$D(d_x, d_y) = \sqrt[p]{(|x_1 - f_{y,1}|)^p + \dots + (|x_j - f_{y,j}|)^p} \quad (2.3)$$