**EKSPLORASI DATA PENGELOMPOKKAN BINTANG NASA : FAKTOR DAN KETERKAITAN ANTAR VARIABEL**

Intan Citra Phonskaningtyas (06211940000007)

Jovanka Alvira Wijaya(06211940000079)

Departemen Statistika ITS/ Analisis Eksplorasi Data

Institut Teknologi Sepuluh Nopember

Email : intancitra@gmail.com, [jovankaawijaya10@gmail.com](mailto:jovankaawijaya10@gmail.com)

**I. PENDAHULUAN**

**1.1 Latar belakang**

Bintang terbentuk dari kumpulan awan dan debu angkasa yang bersatu selama lebih dari satu tahun cahaya. Gravitasi adalah faktor yang menarik debu-debu angkasa untuk bersatu dan menjadi bintang dengan bentuk bola yang sangat besar, hasil reaksi tersebut membuat inti bintang begitu panas hingga dapat mencapai 15 juta derajat. Bintang baru memiliki variasi ukuran dan warna, dengan jangkauan warna biru hingga merah berdasarkan suhu yang dimiliki bintang tersebut. NASA (*National Aeronautics and Space Administration*) mengklasifikasikan bintang yang ada di alam semesta menjadi beberapa tipe, 5 diantaranya yakni *red dwarf, brown dwarf, white dwarf, main sequence, super giants,*dan *hyper giants*. *Red dwarf* adalah tipe bintang yang memiliki temperatur dan ukuran terkecil jika dibandingkan dengan bintang tipe lainnya. Saat bintang ini mati karena kehabisan bahan bakar yang dihasilkan oleh bintang itu sendiri, *red dwarf* akan berubah menjadi bintang tipe *white dwarf*. *Brown dwarf* biasa disebut oleh para saintis *dengan the failed star*s atau bintang gagal, adalah bintang yang memiliki ukurang kecil sehingga tidak seperti bintang lainnya bintang ini terkadang

tidak memiliki cukup fusi hidrogen. *White dwarf* adalah bintang yang telah kehabisan bahan nuklir dari inti bintang sehingga *white dwarf* dapat dikatakan sebagai jenis yang muncul ketika bintang lainnya mulai kehabisan bahan bakar atau dalam tahapan menjadi bintang mati. *Main sequence stars* adalah tipe bintang yang paling bannyak ditemukan saat ini, disebut *main sequence* karena lokasi spesifik pusat massa bintang terletak pada *main sequence* nya. *Super giants* dan *hyper giants* secara garis besar dibedakan melalui ukuran dan umur bintang, dimana *hyper giant*s akan memiliki umur dan ukurang yang lebih besar dibandingkan *supergiants*. Ada banyak faktor lain yang disebutkan memiliki andil dalam pengelompokkan bintang-bintang menurut NASA seperti radius, magnitude, warna, dan lainnya. Salah satu tujuan dibentuknya pengklasifikasi bintang adalah agar manusia dapat membedakan ciri bintang satu dan lainnya, selain itu untuk menunjang kemudahan para saintis untuk melakukan penelitian lebih lanjut.

Dalam karya kali ini, akan dilakukan eksplorasi data mengenai pengaruh faktor-faktor yang digunakan sebagai pengklasifikasi tipe bintang dan hubungan apa yang dimiliki antar faktor tersebut, karena hubungan ini sangat penting untuk mengetahui apakah ada keterkaitan antar faktor dan bagaimana hal tersebut dapat membuat suatu bintang diklasifikasikan dalam tipe tertentu.

**1.2 Rumusan masalah**

Rumusan masalah dalam karya kali ini adalah bagaimana eksplorasi data dari faktor-faktor pengklasifikasi tipe bintang dan bagaimana hubungan yang dimiliki antar variabelnya?

**1.3 Tujuan**

Tujuan dibuatnya karya ini adalah untuk mengetahui eksplorasi data dari faktor-faktor pengklasifikasi tipe bintang dan bagaimana hubungan yang dimiliki antar variabelnya

|  |  |  |
| --- | --- | --- |
| Variabel | Tipe Data | Deskripsi |
| *Star Type* | Nominal | Tipe bintang (Red Dwarfs - 0,Brown Dwarfs - 1, White Dwarfs - 2, Main Sequence - 3, Super Giants - 4, Hyper Giants - 5) |
| *Temperature* | Numerik | Suhu bintang (1939˚K - 40000˚K) |
| *Relative Luminousity* | Numerik | Total energi yang dipancarkan suatu bintang ke segala arah per detik (8e-05 L☉ - 849420 L☉) |
| *Relative Radious* | Numerik | Radius bintang (0.0084 R⊙ - 1984.5 R⊙) |
| *Absolute Magnitude* | Numerik | Kecerahan bintang dengan keterangan semakin kecil magnitude, semakin terang bintang tersebut (-11.92 Mw - 20.06 Mw) |
| *Color* | Nominal | Warna bintang, semakin merah suatu bintang maka suhunya semakin tinggi (17 Unique Values) |
| *Spectral Class* | Nominal | Klasifikasi kelas spektrum bintang (M,B, A,F,O,K,G) |

1. **METODOLOGI**
   1. **Sumber data**

Data yang digunakan dalam karya kali ini adalah data Star Type Classification/NASA oleh Baris Dincer yang diambil dari platform repository data Kaggle. Data yang digunakan dilakukan pengambilan pada :

Hari/Tanggal : Rabu, 23 Juni 2021

Pukul : 08.00

* 1. **Variabel Penelitian**

Variabel yang digunakan berjumlah 7 variabel dengan keterangan sebagai berikut.

**2.3. Langkah Analisis**

Analisis yang dilakukan dalam eksplorasi data kali ini adalah sebagai berikut.

* Merumuskan masalah
* Melakukan studi literatur
* Mencari data yang sesuai
* Melakukan eksplorasi data dengan terlebih dahulu membersihkan data dan menyiap-kan data sebelum dilakukannya eksplorasi lebih lanjut
* Melakukan analisis dari hasil eksplorasi yang dilakukan
* Melakukan penarikan kesimpulan