

BAB V

KESIMPULAN

5.1 KESIMPULAN

Untuk kesimpulan yang dapat diambil dari hasil penelitian yang telah dilakukan pada bab-bab sebelumnya, yaitu :

1. Penelitian ini menggunakan *dataset online* yang bersumber dari [www.kaggle.com](https://www.kaggle.com/datasets/imohtn/video-games-rating-by-esrb) (<https://www.kaggle.com/datasets/imohtn/video-games-rating-by-esrb>), berjumlah 1.895 data yang belum di-*cleaning* dan 1.222 data yang sudah di-*cleaning*. *Dataset* ini memiliki 33 atribut termasuk atribut *class* yang akan menjadi hasil keputusan yang membentuk sebuah pohon (*decision tree*).
2. Dikarenakan banyaknya atribut dari *dataset* tersebut, dilakukan proses seleksi fitur sehingga atribut tersebut menjadi 28 dari 32 atribut. Berikut merupakan atribut yang berpengaruh terhadap klasifikasi *games rating* :
Alcohol Reference, Animated Blood, Blood, Blood and Gore, Cartoon Violence, Crude Humor, Drug Reference, Fantasy Violence, Intense Violence, Language, Mature Humor, Mild Blood, Mild Cartoon Violence, Mild Fantasy Violence, Mild Language, Mild Suggestive Themes, Mild Violence, Nudity, Partial Nudity, Sexual Content, Sexual Themes, Simulated Gambling, Strong Language, Strong Sexual Content, Suggestive Themes, Use of Alcohol, Use of Drugs and Alcohol, Violence.

3. Hasil klasifikasi menggunakan *Data Training* sebagai model pohon keputusan menghasilkan bentuk pohon dan jumlah *node* yang berbeda, seperti :
 - a.) Menggunakan *Data Training* 60% Pada RapidMiner = 28 node.
 - b.) Menggunakan *Data Training* 70% Pada RapidMiner = 25 node.
 - c.) Menggunakan *Data Training* 80% Pada RapidMiner = 23 node.
4. Hasil klasifikasi menggunakan *Data Testing* untuk membandingkan akurasi terbesar yang dihasilkan dari 2 metode dan 2 model pengujian :
 - a.) Hasil dari *Algoritma C4.5 (Use Training Set)*
Accuracy = 86.48% pada *Testing* 20%
 - b.) Hasil dari *Algoritma C4.5 (10 Fold-Cross Validation)*
Accuracy = 80.50% pada *Testing* 40%
 - c.) Hasil dari *Algoritma Naïve Bayes (Use Training Set)*
Accuracy = 78.28% pada *Testing* 20%
 - d.) Hasil dari *Algoritma Naïve Bayes (10 Fold-Cross Validation)*
Accuracy = 75.00% pada *Testing* 20%

Dapat disimpulkan bahwa metode terbaik yang memperoleh hasil akurasi terbesar menggunakan *Use Training Set* dan *10 Fold-Cross Validation* adalah metode *Algoritma C4.5*.

5.2 SARAN

Adapun saran yang dapat penulis berikan dari hasil penelitian ini adalah sebagai berikut :

1. Diharapkan kedepannya penelitian ini menggunakan *dataset games rating* lebih banyak lagi serta mencakup atribut lainnya agar mendapatkan hasil akurasi yang optimal dan tinggi.
2. Diharapkan data ini dapat diujikan menggunakan metode dan algoritma *data mining* lainnya.
3. Diharapkan kedepannya adanya penelitian yang melakukan perbandingan dengan metode dan algoritma lainnya.