

Penerapan Algoritma *Decision Tree Classifier* dalam menentukan kualitas Wine

(Desember 2021)

Widodo, Akbar Maulana, Qisra Lutfi Ranev.

Institut Teknologi Sumatera

Institut Teknologi Sumatera

Institut Teknologi Sumatera

Email: widodo.119140156@student.itera.ac.id, akbar.119140151@student.itera.ac.id, qisra.14116140@student.itera.ac.id

ABSTRAK Teknologi berkembang sangat cepat di masa kini, dimana artificial intelligence menjadi pusat tren dan juga sudah berkembang secara signifikan. Machine Learning merupakan salah satu cabang dari artificial intelligence yang semakin banyak penggunaannya, salah satu jenis algoritma dalam machine learning adalah supervised learning, dimana proses pembelajarannya “diawasi”. Supervised learning membutuhkan data yang telah diberi label untuk membangun sebuah model. Decision Tree Classifier merupakan algoritma supervised learning yang membangun pohon node dalam pengklasifikasian dataset. pada kesempatan kali ini penulis ingin menggunakan algoritma tersebut untuk mengecek kualitas dari wine menggunakan dataset red wine yang telah disediakan.

INDEX TERMS Artificial Intelligence, Supervised Learning, Machine Learning, Kualitas Wine, Decision Tree Classifier

I. INTRODUCTION

Wine merupakan minuman yang memiliki sejarah hingga sekitar tahun 6000 SM. Wine berasal dari mesopotamia dan menyebar ke seluruh dunia. Wine diciptakan secara tidak sengaja oleh seorang wanita yang memetik buah anggur dari ladang miliknya.

Dengan berkembangnya zaman, banyak teknologi yang diciptakan untuk mempermudah manusia. Salah satunya adalah machine learning yang dapat digunakan dalam menentukan suatu hal

secara otomatis. Machine learning merupakan salah satu cabang dari bidang kecerdasan buatan yang berkembang pesat saat ini. Machine learning merupakan mesin yang dapat belajar dan kemudian mengambil keputusan.

Ada beberapa metode yang dapat digunakan dalam penerapan machine learning, salah satunya adalah Decision Tree Classifier. Algoritma Decision Tree Classifier cukup populer untuk digunakan dikarenakan mudah untuk diinterpretasikan oleh manusia.

II. STUDI LITERATUR

2.1 Artificial intelligence

Artificial intelligence atau yang dalam bahasa Indonesia yaitu kecerdasan buatan merupakan ilmu komputer yang mempelajari mesin agar dapat menyelesaikan suatu pekerjaan layaknya manusia. Dengan adanya kecerdasan buatan diharapkan manusia dapat menciptakan mesin yang dapat berpikir dan bertindak tanpa masukan manusia. Salah satu program AI yang cukup terkenal adalah komputer Deep Blue yang dapat mengalahkan juara catur dunia pada tahun 1997.

2.2 Machine Learning

Machine learning merupakan salah satu cabang ilmu pengetahuan yang merupakan bagian dari artificial intelligence. Seperti namanya, machine learning adalah mesin yang dapat belajar dengan sendirinya. Machine learning berusaha untuk menirukan cara manusia berpikir. Machine learning umumnya dapat dibagi menjadi 2 yaitu klasifikasi dan prediksi. Klasifikasi merupakan proses machine learning untuk memilah suatu data berdasarkan ciri tertentu dan prediksi adalah mesin menebak keluaran berdasarkan masukan yang diberikan.

2.3 Supervised learning

Pada penerapan machine learning terdapat beberapa teknik yang dapat digunakan salah satunya ialah supervised learning. Supervised learning merupakan teknik yang menggunakan dataset berlabel untuk melakukan pembelajaran kepada mesin. Sehingga mesin dapat mengidentifikasi label input berdasarkan dataset yang dimiliki. Algoritma yang termasuk dalam supervised learning adalah Decision Tree, K-Nearest Neighbor (KNN), Naive Bayes, Regresi, dan Support Vector Machine.

2.4 Pengelompokan (Clustering)

Pengelompokan atau yang lebih sering disebut dengan clustering merupakan proses pembagian objek menjadi kelompok-kelompok yang serupa. Dalam machine learning pengelompokan dilakukan berdasarkan ciri-ciri dari data yang ada. Terkadang pada pengelompokan machine learning terdapat atribut-atribut yang tidak diperlukan. Dalam kasus tersebut atribut yang tidak diperlukan dapat dihapus untuk meningkatkan performa dari machine learning tersebut.

2.5 Algoritma Decision Tree Classifier

Algoritma Decision Tree Classifier merupakan teknik yang dapat digunakan dalam

machine learning. Algoritma tersebut menggunakan aturan klasifikasi yang berstruktur dan menjadikan dataset beberapa bagian dan dilakukan secara berulang. Seperti namanya algoritma ini memiliki struktur seperti pohon yang terdiri dari node-node data.

III. METODE

Dalam penelitian ini kami menggunakan metode Decision Tree Classifier untuk menentukan kualitas wine. Metode ini kami pilih dikarenakan mudah dipahami dan tidak memerlukan banyak persiapan data. Metode ini sangat cocok untuk menentukan klasifikasi sebuah data. Pada metode ini terdapat decision node yang terbuat dari atribut dan memecah dataset menjadi node yang lebih kecil. Hal ini dilakukan secara berulang kali sehingga jika digambarkan akan menjadi seperti sebuah pohon. Algoritma yang kami buat dijalankan menggunakan Google Colab. Kami menggunakan Google Colab dikarenakan Google Colab memiliki banyak library yang mendukung dalam pembuatan machine learning.

IV. DATA SET

Dataset yang kami gunakan dalam penelitian ini adalah winequality-red.csv. Dataset tersebut kami pilih karena menurut kami dataset tersebut memiliki parameter penentu kualitas yang cukup banyak. Sehingga algoritma yang kami buat dapat menjadi lebih akurat. Pada eksperimen ini kami melakukan binerisasi pada nilai kualitas wine. Dimana, kualitas 3-6 menjadi 0 yang artinya bad quality dan kualitas 7-8 menjadi 1 yang artinya good quality. Hal ini dilakukan untuk mempermudah pengguna dalam mengetahui kualitas wine yang sudah dihasilkan. Dataset yang kami miliki akan kami bagi menjadi 2 yaitu dataset training dan dataset testing. Dataset training berisikan 80% dari total dataset yang ada, dan dataset testing berisikan 20% dari total dataset yang ada.

V. HASIL EKSPERIMEN & PEMBAHASAN

Dari eksperimen yang telah dilakukan dengan menggunakan algoritma Decision Tree Classifier didapatkan data tree seperti berikut.

```
X_10 <= 10.9 ? 0.030622084571308328
left:X_9 <= 0.67 ? 0.086176747016306883
left:X_3 <= 1.2 ? 0.0014071898164856739
left:0.0
right:X_4 <= 0.062 ? 0.00097250736429556744
left:0.0
right:0.0
right:X_0 <= 11.5 ? 0.028601226602465557
left:X_4 <= 0.061 ? 0.021458246370039513
left:0.0
right:0.0
right:X_1 <= 0.33 ? 0.1601307189542484
left:1.0
right:0.0
right:X_9 <= 0.68 ? 0.04517501259105927
left:X_1 <= 0.35 ? 0.03626687868130773
left:X_3 <= 3.6 ? 0.06274946625823813
left:0.0
right:0.0
right:X_2 <= 0.73 ? 0.01984498703774465
left:1.0
right:0.0
right:X_10 <= 11.5 ? 0.06400099129083732
left:X_1 <= 0.37 ? 0.06471998001505018
left:1.0
right:0.0
right:X_5 <= 18.0 ? 0.0573099941520467866
left:1.0
right:0.0
```

[illegible]

1. 90% - 100% = Excellent Classification
2. 80% - 90% = Good Classification
3. 70% - 80% = Fair Classification
4. 60% - 70% = Poor Classification
5. 50% - 60% = Failure

MENGHITUNG AKURASI PENGETESAN

0.88125

PENGETESAN MANUAL

[1.0]
Good Quality

total acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	ph	alcohols	alcohol quality	
7.4	0.70	0.04	1.9	0.076	11.0	34.0	0.98780	3.51	0.56	9.4	5
7.8	0.88	0.00	2.6	0.088	25.0	67.0	0.98680	3.28	0.68	9.8	5
7.8	0.70	0.04	2.3	0.092	19.0	54.0	0.98700	3.56	0.68	9.8	5
7.2	0.56	0.01	0.073	0.077	17.0	60.0	0.98620	3.16	0.58	9.6	5
7.4	0.70	0.00	1.9	0.076	11.0	34.0	0.98780	3.51	0.56	9.4	5
7.9	0.37	0.23	1.8	0.077	23.0	49.0	0.98630	3.28	0.67	9.3	5
12.5	0.56	0.49	2.4	0.064	5.0	27.0	0.99990	3.96	0.67	10.9	5
11.6	0.36	0.92	1.3	0.076	6.0	16.0	0.99860	3.30	0.72	10.2	7
8.1	0.87	0.01	0.3	0.091	26.0	66.0	1.00020	3.34	0.71	9.8	6
7.9	0.35	0.46	3.5	0.078	15.0	27.0	0.98730	3.35	0.66	12.8	8

Dari eksperimen yang telah dilakukan dapat disimpulkan, bahwa eksperimen dengan menggunakan algoritma Decision Tree Classifier pada dataset wine ini berhasil dengan mendapat akurasi 88% dan predikat “Good Classifier”, juga setelah dilakukan beberapa pengecekan klasifikasi secara manual sebagian besar data klasifikasi yang dihasilkan benar. Selanjutnya, menurut kami apabila membutuhkan akurasi yang lebih tinggi dapat menggunakan kedalaman node yang lebih besar.

Refrensi

- [1] S. Rais, "Persepsi wisatawan asing terhadap produk hatten wine sebagai seni kuliner bali untuk daya tarik wisata," *J. Account. Manag. Innov.*, vol. 1, no. 2, pp. 159–177, 2017, [Online]. Available: <https://ejournal.medan.uph.edu/index.php/jam/article/download/149/39>.
- [2] Iykra, "Mengenal Decision Tree dan Manfaatnya," *medium.com*, 2018. <https://medium.com/iykra/mengenal-decision-tree-dan-manfaatnya-b98cf3cf6a8d> (accessed Dec. 24, 2021).
- [3] Syafnidawaty, "Kecerdasan Buatan," *raharja.ac.id*, 2011. <https://raharja.ac.id/2020/11/27/kecerdasan-buatan/> (accessed Dec. 24, 2021).
- [4] A. Sudibyo, T. Asra, and B. Rifai, "Klasifikasi Seleksi Atribut Pada Serangan Spam Menggunakan Metode Algoritma Decision Tree," *J. Pilar Nusa Mandiri*, vol. 14, no. 2, p. 145, 2018, doi: 10.33480/pilar.v14i2.874.
- [5] A. Ahmad Hania, "Mengenal Artificial Intelligence, Machine Learning, & Deep Learning," *J. Teknol. Indones.*, no. June, 2017, [Online]. Available: <https://amt-it.com/mengenal-perbedaan-artificial-intelligence-machine-learning-deep-learning/>.
- [6] E. Retnoningsih and R. Pramudita, "Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python," *Bina Insa. Ict J.*, vol. 7, no. 2, p. 156, 2020, doi: 10.51211/biict.v7i2.1422.
- [7] I. Setiawati, A. Permana, and A. Hermawan, "Implementasi Decision Tree Untuk Mendiagnosis Penyakit Liver," *J. Inf. Syst. Manag.*, vol. 1, no. 1, pp. 13–17, 2019, doi: 10.24076/joism.2019v1i1.17.