

Implementasi Algoritma Naïve Bayes Untuk Klasifikasi Kesegaran Buah jeruk

Guntur Firmansyah^{1*}, Arief Hermawan²

^{1,2} Universitas Teknologi Yogyakarta, D.I Yogyakarta, Indonesia

e-mail: ¹guntur.5200411138@student.uty.ac.id, ²ariefdb@uty.ac.id

Informasi Artikel

Diterima: 14-06-2023

Direvisi: 04-10-2023

Disetujui: 05-10-2023

Abstrak

Jeruk adalah salah satu buah yang paling populer di dunia. Jeruk merupakan buah sitrus yang memiliki banyak kandungan vitamin C dan anti oksidan, selain itu jeruk juga bermanfaat bagi kekebalan tubuh. Buah jeruk banyak tumbuh di daerah tropis salah satunya Indonesia. Produktivitas buah jeruk di Indonesia sangat tinggi dengan hal tersebut maka proses pemilahan buah jeruk berkualitas akan membutuhkan waktu yang lama apabila dilakukan secara manual. Hal ini tentunya akan mengurangi efektivitas dan efisiensi waktu yang ada. Sehingga dibutuhkannya suatu sistem komputasi dengan suatu algoritma untuk menunjang permasalahan tersebut. Perancangan sistem ini akan menggunakan algoritma naïve bayes untuk mengklasifikasikan kualitas buah jeruk. Sistem ini akan mempermudah proses pemilahan jeruk berkualitas dan manajemen waktu akan lebih efektif. Hasil akhir dari sistem ini mendapatkan hasil akurasi pelatihan 72,7% dan pengujian 62,6%.

Kata Kunci: *Naïve Bayes, Klasifikasi, Kesegaran Buah jeruk*

Abstract

Oranges are one of the most popular fruits in the world. Oranges are citrus fruits that contain lots of vitamin C and anti-oxidants, besides that oranges are also beneficial for the immune system. Citrus fruits grow a lot in tropical areas, one of which is Indonesia. The productivity of citrus fruits in Indonesia is very high, so that the process of sorting quality citrus fruits will take a long time if done manually. This of course will reduce effectiveness and efficiency existing time. So we need a processing system with an algorithm to support these problems. The design of this system will use the naïve Bayes algorithm to classify the quality of citrus fruits. This system will simplify the process of sorting quality oranges and more effective time management. The final result of this system is to get 72,7% training accuracy and 62,6% testing.

Keywords: *Naïve Bayes, Classification, Oranges Refreshment*

1. Pendahuluan

Jeruk adalah buah musiman yang berasal dari Asia. Terdapat beberapa keyakinan bahwa Cina merupakan tempat asal dari buah jeruk tumbuh. Namun setelah ratusan tahun jeruk juga tumbuh di Indonesia secara alami maupun dibudidayakan (Yanto, B., et al., 2021). Selain rasa buahnya yang manis ternyata kulit jeruk juga dapat dimanfaatkan sebagai obat jerawat (Michiko., et al., 2020)

Kematangan buah jeruk dapat dilihat dari warna serta teksturnya. Jeruk dalam keadaan matang biasanya memiliki tekstur kulit yang lebih halus dan tipis. Dengan adanya kemajuan teknologi saat ini sangat memungkinkan untuk membuat suatu sistem untuk mengklasifikasikan buah jeruk. Karena proses penyortiran secara manual cukup terbatas secara fisik atau kondisi tertentu yang dapat mengurangi tingkat efektivitas dan efisiensi (Haba & Pelangi, 2020).

Oleh karena itu diharapkan sistem ini dapat membantu agar terjaganya konsistensi dari penyortiran dengan cara mengklasifikasikan citra.

Citra merupakan suatu matrik dua dimensi dari fungsi intensitas cahaya (Mandala., et al., 2021). Secara umum tahapan klasifikasi citra digital melalui beberapa tahapan diantaranya yaitu akuisisi citra, pra pengolahan citra, ekstraksi ciri, pelatihan, pengujian dan pengukuran akurasi.

Algoritma klasifikasi ada beberapa metode, salah satunya adalah Naïve Bayes. Naïve Bayes disebut juga multinomial yang merupakan metode klasifikasi yang dapat memprediksi dengan cara memanfaatkan probabilitas dan statistic (Wibisono., et al., 2020). Naïve bayes termasuk kedalam kategori machine learning. Machine learning dilatih melalui sekumpulan data historis (Hayami., et



al., 2022). Machine learning bersifat prediktif dan deskriptif yang bertujuan untuk memprediksi berdasarkan data dan mendapatkan informasi dari data (Oikonomidis., et al., 2022).

Pada penelitian sebelumnya, Naïve Bayes digunakan untuk mengklasifikasikan tingkat kematangan buah pisang Kepok. Pengklasifikasian tingkat kematangannya dibagi dari yang masih hijau, agak kekuningan, baru ujungnya yang menguning, hingga yang telah terlalu masak. Tingkat akurasi dari penelitian ini mencapai 81% (Hakim., et al., 2022).

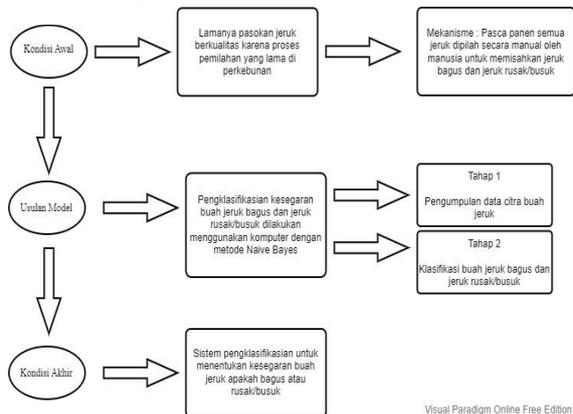
Adapun penelitian sebelumnya yang memanfaatkan Naïve Bayes untuk klasifikasi penerima program keluarga harapan. Nilai akurasi yang didapatkan mencapai 88% (Arifin., et al., 2022).

Berdaskan penelitian sebelumnya maka klasifikasi kesegaran buah jeruk dapat menggunakan metode naïve bayes. Penelitian ini bertujuan untuk mengklasifikasikan kesegaran buah jeruk yang memiliki akurasi akurat. Serta akan dibangun suatu sistem menggunakan bahasa *Python* dan *GUI Tkinter*.

2. Metode Penelitian

2.1. Kerangka Penelitian

Kerangka penelitian pada umumnya adalah tahapan konsep yang digunakan dalam merancang sistem. Kerangka penelitian yang digunakan pada penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Kerangka Penelitian

Kondisi awal dari permasalahan dalam penelitian ini yaitu lamanya proses pemilahan yang dilakukan secara manual oleh manusia. Selain membutuhkan waktu yang lama, pemilahan secara manual juga terpengaruh oleh kemampuan fisik seseorang. Dengan adanya kondisi tersebut maka akan dapat mempengaruhi tingkat efektivitas dan efisiensi dari proses pemilahan.

Dengan permasalahan kondisi awal tersebut maka dibutuhkan solusi untuk menyelesaikannya. Oleh karena itu pada penelitian ini mengusulkan suatu sistem

komputer yang dapat melakukan klasifikasi kesegaran buah jeruk dengan menggunakan algoritma naïve bayes.

Kondisi akhir dari usulan model ini adalah terciptanya sistem komputer yang dapat melakukan pengklasifikasian terhadap jeruk panen dengan melakukan *image classification*. Dengan terbangunnya sistem tersebut diharapkan dapat mengatasi permasalahan pada kondisi awal.

2.2. Pengumpulan Data

Dalam tahap pengumpulan data dilakukan dengan melakukan observasi. Observasi merupakan salah satu metode pengumpulan data dalam penelitian kualitatif. Dalam penelitian kualitatif, data akan menjadi lebih baik dan lebih valid ketika dilengkapi dengan analisa (Rachmawati, 2017). Data yang terkumpul berupa data citra atau data gambar yang terbagi menjadi 2 kelas yaitu jeruk bagus dan jeruk rusak/busuk. Data diperoleh dengan cara melakukan pengambilan gambar terhadap jeruk secara pribadi dan pelabelan dilakukan secara manual dengan bantuan ahli buah. Dari sebanyak 133 data citra yang didapatkan terdiri dari 79 jeruk bagus dan 54 jeruk rusak/busuk. Sampel data citra yang diperoleh dapat dilihat pada gambar 2.

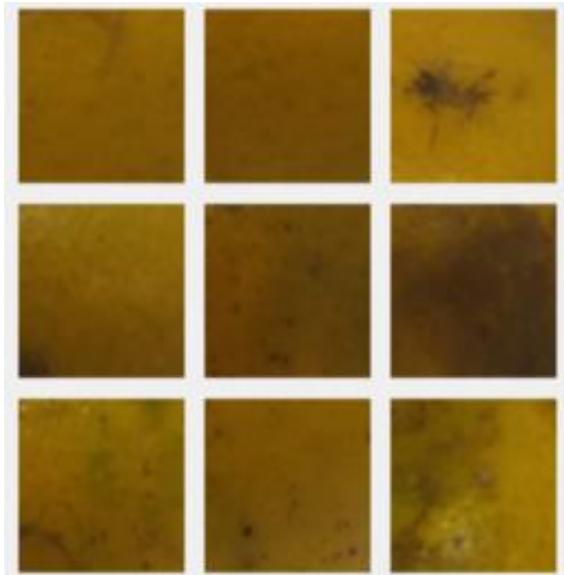


Gambar 2. Sampel Data

2.3 Preprocessing Data

Preprocessing data akan mengolah data mentah untuk di *resize* dan *cropping*. *Cropping* digunakan untuk membuat gambar lebih fokus serta seragam (Peryanto., et al., 2020). Serta akan dihitung nilai rata-rata dari tiap layer gambar RGB. Dataset train digunakan untuk mengevaluasi model (Stephen, 2019). Hasil *cropping* dapat dilihat pada gambar 3, serta hasil

perhitungan mean rgb dapat dilihat pada gambar 4.



Gambar 3. Cropping

	A
1	R,G,B,Label
2	135.5046216984352,140.66987962087796,140.41481158763915,Bagus
3	149.9065035247322,153.48900982592943,155.2919472668557,Bagus
4	134.13420155298255,138.846030081653,138.7911489511132,Bagus
5	133.95517699144088,139.79528903460405,140.3509282779353,Bagus
6	132.9169095253098,139.7818834015963,140.17322220909472,Bagus
7	142.7941165852762,148.10471309861373,150.25214716577398,Bagus
8	147.30993767722117,153.3856595253098,155.42999025283552,Bagus
9	104.8367940689981,110.15675212665406,114.18144740469438,Bagus

Gambar 4. Mean RGB

Citra digital direpresentasikan dalam bentuk matriks piksel. Setiap piksel dalam gambar warna memiliki tiga komponen warna: Red (R), Green (G), dan Blue (B), dan setiap komponen tersebut direpresentasikan sebagai nilai integer antara 0 hingga 255 (8-bit per komponen). Nilai mean dari tiap layer akan digunakan sebagai acuan untuk klasifikasi kesegaran citra jeruk. Pada tabel 1 dapat dilihat bahwa kelas rusak/busuk memiliki nilai lebih tinggi dibandingkan kelas bagus.

Tabel 1. Nilai mean

Kelas	Mean		
	Red	Green	Blue
Bagus	133.813	138.328	138.918
Rusak/ Busuk	154.233	156.617	158.807

2.4. Naïve Bayes

Naïve bayes adalah sebuah algoritma supervised learning berdasarkan teorema Bayes yang digunakan untuk memecahkan masalah klasifikasi dengan mengikuti pendekatan

probabilitas (Jumeilah, 2018). Naïve Bayes merupakan algoritma yang akan menghitung proses *Training* dan akan melakukan Prediksi terhadap data Uji. Pengujian dilakukan untuk mengklasifikasikan data yang kelasnya tidak diketahui (Adi, 2018). Naïve Bayes merupakan algoritma klasik yang memiliki algoritma sederhana dan efisiensi komputasi yang tinggi. Naïve Bayes hanya perlu memperkirakan parameter yang diperlukan (Chen., et al., 2021). Dasar teorema Naïve Bayes yaitu sebagai berikut :

$$P(C_i | X) = \frac{P(X|C_i).P(C_i)}{P(X)} \quad (1)$$

Dengan keterangan sebagai berikut :

- X = Data dengan class yang belum diketahui
- C_i = Hipotesis data X merupakan suatu class spesifik
- P(C_i|X) = Probabilitas label kelas C_i dengan kriteria masukan X
- P(X|C_i) = Probabilitas kriteria masukan X dengan label kelas C_i
- P(C_i) = Probabilitas label kelas C_i
- P(X) = Probabilitas label kelas X

3. Hasil dan Pembahasan

3.1. Hasil Pelatihan

Dalam melakukan pelatihan peneliti melakukan beberapa percobaan dengan melakukan perubahan terhadap rasio pembagian data. Hasil akurasi tertinggi didapatkan dengan rasio pembagian data 50% : 50%. Berikut hasil dari setiap percobaan rasio pembagian data dapat dilihat pada tabel 1.

Tabel 2. Hasil Pelatihan

No	Data Latih	Data Uji	Akurasi Training	Akurasi Testing
1	80%	20%	71,%	59%
2	70%	30%	70%	60%
3	60%	40%	72,5%	61%
4	50%	50%	72,7%	62,6%

3.2. Implementasi GUI

1) Halaman Utama

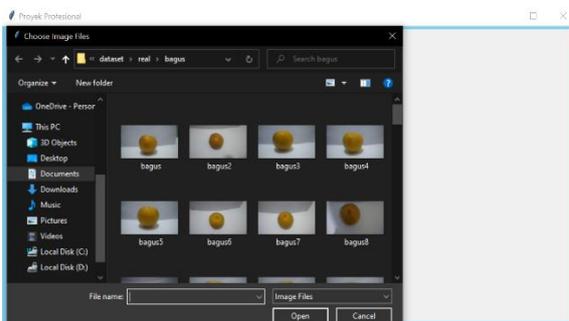
Halaman utama merupakan tampilan awal sistem, terdapat 9 tombol yang berfungsi diantaranya memilih data, mengkonversi data, menggabungkan *file* csv, melakukan pelatihan, melakukan pengujian serta menghapus gambar yang telah ditampilkan. Halaman utama dapat dilihat pada gambar 5.



Gambar 5. Halaman Utama

2) Pilih Data

Pada gambar 6 merupakan tampilan pada menu pilih data yang akan digunakan sebagai data latih. Serta gambar 7 adalah tampilan dari data yang telah dipilih.



Gambar 6. Tampilan Pilih Data



Gambar 7 Tampilan Data Yang Dipilih

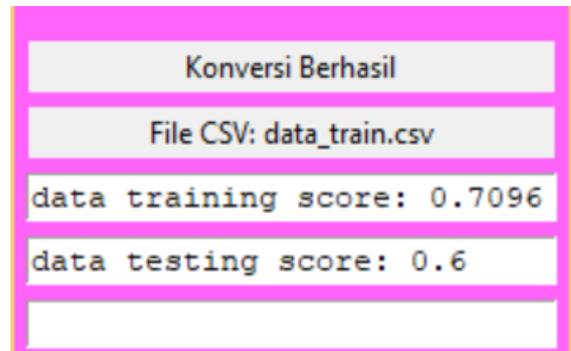
3) Pelatihan

Pada gambar 8 merupakan tampilan apabila *mean rgb* telah dihitung dan telah berhasil dikonversi menjadi *file csv*.



Gambar 8 Tampilan Konversi Berhasil

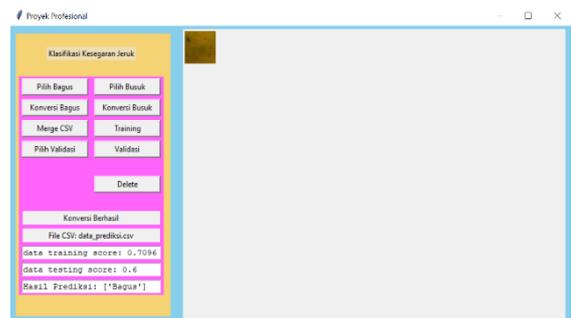
Pada gambar 9 merupakan tampilan hasil dari proses pelatihan yang menampilkan akurasi yang didapat.



Gambar 9. Tampilan Hasil Akurasi

4) Pengujian

Setelah mendapatkan model terbaik, model dapat digunakan untuk memprediksi data baru. Hasil prediksi data baru dapat dilihat pada gambar 10. Berdasarkan pengujian model terhadap data baru tersebut termasuk kedalam kelas Jeruk Bagus.



Gambar 10. Hasil Prediksi Data Baru

4. Kesimpulan

Kesimpulan dari penelitian Implementasi Algoritma Naïve Bayes Untuk Klasifikasi Kesegaran Jeruk berhasil dilakukan. Sistem berhasil dibangun menggunakan bahasa pemrograman *python*. Penelitian ini berhasil memanfaatkan naïve bayes sebagai algoritma pelatihan sistem. Hasil akurasi training sebesar 72,7% serta akurasi testing sebesar 62,6%. Sistem yang telah dirancang ini dapat dikembangkan misalnya dengan menambahkan fitur baru. Menambahkan proses pengolahan data citra serta menambah jumlah dataset untuk mendapatkan hasil akurasi yang lebih baik.

Referensi

- Arifin, A. A. A., Handoko, W., & Efendi, Z. (2022). Implementasi Metode Naive Bayes Untuk Klasifikasi Penerima Program Keluarga Harapan. *J-Com (Journal of Computer)*, 2(1), 21–26. <https://doi.org/10.33330/j-com.v2i1.1577>

- Chen, H., Hu, S., Hua, R. et al. Improved naive Bayes classification algorithm for traffic risk management. *EURASIP J. Adv. Signal Process.* 2021, 30 (2021).
- F. S. Jumeilah, "Klasifikasi Opini Masyarakat Terhadap Jasa Ekspedisi JNE dengan Naïve Bayes," *J. Sist. Inf. Bisnis*, vol. 8, no. 1, pp. 92–98, 2018
- Haba, A. R. K., & Pelangi, K. C. (2019). Sistem Cerdas Dalam Klasifikasi Kematangan Buah Jeruk Berdasarkan Fitur Ekstraksi Gcm Dengan Metode Naïve Bayes. *Jurnal Teknologi Dan Manajemen Informatika*, 5(2).
<https://doi.org/10.26905/jtmi.v5i2.3935>
- Hakim, Z., Rahayu, S., dan Irawati, K. (2022). Klasifikasi Tingkat Kematangan Buah Pisang Kepok Menggunakan Algoritma Naive Bayes. *Academic Journal of Computer Science Research* 4(1): 8–11.
- Hayami, R., Soni, & Gunawan, I. (2022). Klasifikasi Jamur Menggunakan Algoritma Naïve Bayes. *Jurnal CoSciTech (Computer Science and Information Technology)*, 3(1),28–33.
<https://doi.org/10.37859/coscitech.v3i1.3685>
- Mandala, J. Y., Warsito, A., Lapono, L. A. S., & Louk, A. C. (2021). Rancang Bangun Prototype Sistem Akuisisi Citra Digital Untuk Proses Tomografi. *Jurnal Fisika: Fisika ...*, 6(1), 32–43.
<http://ejournal.undana.ac.id/FISA/article/view/4500>
- Michiko, M., Manalu, C. V., & Mutia, M. S. (2020). Uji Efektivitas Ekstrak Etanol Kulit Jeruk Manis (*Citrus Sinensis*) Terhadap Bakteri *Propionibacterium Acnes*. (*Jurnal Ilmiah Mahasiswa Kesehatan Masyarakat*), 5(1), 6–9.
<https://doi.org/10.37887/jimkesmas.v5i1.10552>
- Oikonomidis, A., Catal, C., & Kassahun, A. (2023). Deep learning for crop yield prediction: a systematic literature review. *New Zealand Journal of Crop and Horticultural Science*, 51(1), 1–26.
<https://doi.org/10.1080/01140671.2022.2032213>
- Peryanto, A., Yudhana, A., & Umar, R. (2020). Klasifikasi Citra Menggunakan Convolutional Neural Network dan K Fold Cross Validation. *Journal of Applied Informatics and Computing*, 4(1), 45–51.
<https://doi.org/10.30871/jaic.v4i1.2017>
- Rachmawati, T. (2017). Metode pengumpulan data dalam penelitian kualitatif. *UNPAR Press*, 1, 1-29.
- S. Adi, "Implementasi Algoritma Naïve Bayes Classifier Untuk Klasifikasi Penerima Beasiswa PPA Di Universitas Amikom Yogyakarta," *J. Mantik Penusa*, vol. 22, no. 1, pp. 11–16, 2018, [Online]. Available: <https://ejournal.pelitanusantara.ac.id/index.php/mantik/art>
- Stephen, S. , R. R. , & S. H. (2019). CNN Clasification Apps. *Jurnal Sistem Informasi Dan Telematika (Telekomunikasi, Multimedia Dan Informatika)*, 10(2).
<http://dx.doi.org/10.36448/jsit.v10i2.1319>
- Wibisono, A. D., Dadi Rizkiono, S., & Wantoro, A. (2020). Filtering Spam Email Menggunakan Metode Naive Bayes. *Telefortech : Journal of Telematics and Information Technology*, 1(1), 9–17.
<https://doi.org/10.33365/tft.v1i1.685>
- Yanto, B., Fimawahib, L., Supriyanto, A., Hayadi, B. H., & Pratama, R. R. (2021). Klasifikasi Tekstur Kematangan Buah Jeruk Manis Berdasarkan Tingkat Kecerahan Warna dengan Metode Deep Learning Convolutional Neural Network. *Inovtek Polbeng - Seri Informatika*, 6(2), 259. <https://doi.org/10.35314/isi.v6i2.2104>