

KLASIFIKASI KUALITAS MUTU JAGUNG DENGAN MENGUNAKAN METODE *DECISION TREE* PADA DINAS PERTANIAN BOJONEGORO

*Classification of corn quality using the decision tree method at the bojonegoro
agriculture office*

Moch. Lutfi¹⁾

¹⁾Program Studi Teknik Informatika, Universitas Yudharta Pasuruan
Email: moch.lutfi@yudharta.ac.id

ABSTRAK

Jagung merupakan sumber pangan di beberapa daerah di Indonesia, selain sebagai sumber pangan manusia jagung juga dapat sebagai bahan pakan ternak. Klasifikasi mutu kualitas jagung sangat diperlukan untuk menghasilkan hasil produksi yang berkualitas. Kualitas mutu jagung ditentukan oleh *attribut* varietas, panjang, bentuk, warna, rasa, teknik, musim, hama, PH dan mutu yang digunakan sebagai proses klasifikasi kualitas mutu dengan algoritma *decision tree*. Hasil penelitian dibandingkan dengan tiga kriteria *decision tree* yaitu *gain ratio*, *information gain* dan *gini index*. Hasil akurasi menunjukkan bahwa nilai kriteria tertinggi sebesar 72,76% pada kriteria *gini index* dan nilai terendah pada kriteria *gain ratio* sebesar 60,67%.

Kata Kunci: klasifikasi kualitas, mutu jagung, metode *decision tree*

ABSTRACT

Corn is a food source in several regions in Indonesia, in addition to being a human food source, corn can also be used as animal feed ingredients. Quality classification of corn is needed to produce quality products. The quality of corn quality is determined by the attributes of varietas, lengths, shapes, colors, flavors, techniques, seasons, pests, PH and the quality used as a classification process of quality using a decision tree algorithm. The results of the study were compared with three decision tree criteria namely *gain ratio*, *information gain* and *Gini index*. Accuracy results show that the highest criterion value is 72.76% in the Gini index criterion and the lowest value is the criterion *gain ratio* of 60.67%.

Keywords: corn quality, decision tree method, quality classification

PENDAHULUAN

Jagung merupakan sumber pangan di beberapa daerah di Indonesia, selain sebagai sumber pangan manusia jagung juga dapat sebagai bahan pakan ternak (Safuan & Hadini, 2012; Siregar & Nugraha, 2018). Untuk meningkatkan produktivitas jagung, hal yang harus

diperhatikan adalah meningkatkan produktivitas mutu kualitas produksi jagung tersebut.

Namun kenyataannya sampai saat ini yang menjadi permasalahan adalah kurangnya kualitas mutu produksi jagung. Tingginya kerusakan biji jagung di akibatkan oleh kurang baiknya dalam

proses penanganan (Firmansyah dkk., 2007).

Selama ini evaluasi kualitas biji jagung dalam proses pengklasifikasian masih dilakukan dengan cara manual melalui pengamatan secara langsung. Evaluasi mutu jagung dengan cara pengamatan secara langsung memiliki beberapa kelemahan, antara lain membutuhkan proses waktu yang lama sehingga hasil dari proses klasifikasi mutu jagung tidak konsisten (Fitriani dkk., 2017; Kusumaningtyas & Asmara, 2016). Ketidak konsistenan ini disebabkan keterbatasan visual manusia, kelelahan dan adanya perbedaan persepsi tentang kualitas oleh masing-masing pengamat (Bustomi & Dzulfikar, 2014).

Klasifikasi merupakan proses pembelajaran secara terbimbing melalui *training set* berdasarkan data histori (Prabowo dkk., 2018; Yaumi, 2017). Pada penelitian ini digunakan *Training set* dengan menggunakan metode *decision tree* pada dataset kualitas mutu produksi jagung pada dinas pertanian kabupaten Bojonegoro.

Decision tree merupakan algoritma yang digunakan untuk proses klasifikasi maupun prediksi melalui tahap pembentukan pohon keputusan. Pohon keputusan digunakan untuk prosedur

penalaran untuk mendapatkan jawaban dari sebuah masalah (Prasetyo, 2014). Quinlan (1986) memperkenalkan dalam ID3 induksi *decision tree* hanya bisa dilakukan pada fitur yang bertipe kategorikal, sedangkan tipe numerik tidak dapat digunakan. Namun Quinlan melakukan perbaikan metode ID3 menjadi C4.5 yang dapat menangani fitur tipe numerik, melakukan pemotongan, dan penurunan *rule set*. Algoritma C4.5 juga dapat menggunakan kriteria *gain* dalam menentukan fitur untuk pemecahan *node* dari akar.

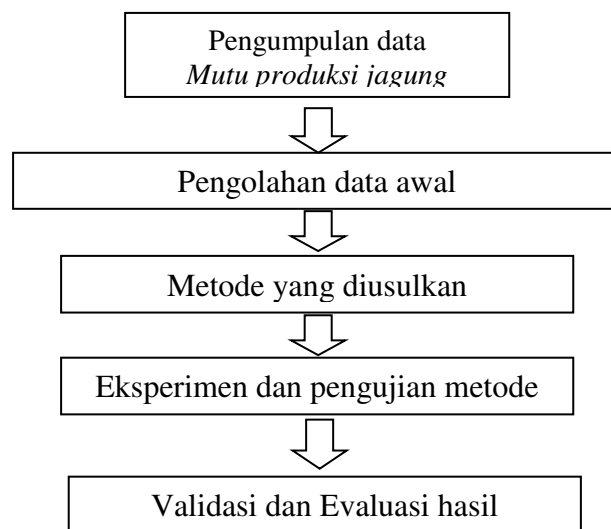
Dalam penelitian ini digunakan metode *decision tree* untuk menentukan kualitas mutu produksi jagung dan melakukan perbandingan akurasi pada kriteria *gain ratio*, *gini indeks*, dan *information gain*.

METODE PELAKSANAAN

Penelitian ini menggunakan pendekatan kuantitatif. Metode kuantitatif bertujuan untuk mencapai pemahaman tentang bagaimana sesuatu dikonstruksi, bagaimana dibangun, atau bagaimana cara kerjanya (Berndtsson dkk., 2007). Penelitian kuantitatif pada umumnya ditekankan oleh hipotesis, yang berdasarkan rumus dan diuji secara bertahap, dengan tujuan untuk

menunjukkan bahwa hipotesisnya benar atau salah. Oleh karena itu, salah satu upaya untuk membenarkan hipotesis, jika hipotesis tahan uji, maka akan dianggap benar sebaliknya jika hipotesisnya di uji tidak tahan maka dianggap salah. Aspek pendekatan kuantitatif adalah untuk menekankan bahwa pengukuran merupakan inti dasarnya karena memberikan hubungan antara observasi dan model, teori, maupun hipotesis (Hamdi & Bahrudin, 2015). Hasil dari penelitian dan metode kuantitatif adalah mengembangkan model, teori, dan hipotesis yang berkaitan dengan fenomena alam.

Metode yang digunakan pada penelitian ini adalah model dari eksperimen. Pada penelitian hasil eksperimen mencakup investigasi antara hubungan sebab-akibat menggunakan pengujian yang dikontrol oleh diri sendiri (Dawson, 2009). Sering kali penelitian semi eksperimental mendapatkan kendala pada tidak cukupnya akses terhadap sampel, masalah etika dan sebagainya. Eksperimen biasanya dilakukan dalam pengembangan, evaluasi dan pemecahan masalah proyek. Karena penelitian yang diakui harus mengikuti tahap prosedur (Dawson, 2009), maka pada penelitian ini dilakukan dengan mengikuti tahapan seperti Gambar 1.



Gambar 1. Tahapan Penelitian

Penjelasan tahap-tahap pada sekedar melihat dan mengambil data Gambar 1. sebagai berikut ini: sampel, tetapi harus mampu

1. Pengumpulan data

Pengumpulan data bukan hanya memiliki kontribusi terhadap

pengetahuan. Data tersebut dapat memberikan penjelasan, hubungan, perbandingan, prediksi, estimasi, generalisasi, dan teori (Dawson, 2009).

2. Pengolahan awal data

Data yang sudah dikumpulkan diolah menggunakan algoritma *softcomputing* untuk mengurangi data yang tidak relevan, atau data dengan atribut yang hilang.

3. Metode yang diusulkan

Untuk menggambarkan alur metode yang diusulkan dan menjelaskan cara kerja metode yang diusulkan maka gambaran penelitian ini digambarkan secara skematik dan disertai dengan rumus dan penghitungan. Sedangkan metode diusulkan akan dibentuk dari data awal yang sudah proses, dan hasil pengolahan metode akan diukur dengan metode yang digunakan.

4. Eksperimen dan pengujian metode

Menjabarkan bagaimana eksperimen yang dilakukan hingga terbentuknya metode, serta menjelaskan cara mengujinya.

5. Validasi dan Evaluasi hasil

Validasi dilakukan dengan mengukur hasil klasifikasi maupun prediksi dibandingkan dengan data awal. Evaluasi hasil dilakukan dengan mengamati hasil klasifikasi maupun

prediksi menggunakan algoritma *soft-computing*. Pengukuran kinerja dilakukan dengan membandingkan nilai *accuracy* dan dibuatkan diagram untuk mengetahui signifikasi perbedaan kinerja metode, sehingga hasil klasifikasi maupun prediksi dari masing-masing algoritma dapat diketahui hasil yang lebih akurat.

Tahapan algoritma *decision tree*

Tahapan membangun pohon keputusan adalah sebagai berikut:

- Siapkan dataset produksi jagung
- Pilih atribut sebagai akar.
- Buat cabang untuk tiap-tiap nilai.
- Bagi kasus pada cabang.
- Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut yang akan digunakan sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut tersebut. Untuk mendapatkan nilai *gain*, terlebih dahulu harus di tentukan nilai Entropinya:

Entropi

Untuk mendapatkan nilai *Information gain* dan *gain Ratio*, terlebih dahulu kita harus menghitung nilai *Entropy* (Madadipouya, 2015).

$$Entropy(S) = \sum_{j=1}^k -p_j * \log_2 p_j$$

Keterangan:

S = himpunan (dataset) kasus

K = banyaknya partisi S

Pj = probabilitas yang di dapat dari sum
(ya) dibagi total kasus.

Sementara untuk perhitungan nilai *gain*
digunakan.

Information gain

Setelah mendapatkan nilai *Entropy*
langkah selanjutnya adalah perhitungan
nilai *Information gain* dari suatu atribut
(Madadipouya, 2015).

$$Gain(A) = Entropy(S) - \sum_{i=1}^k \frac{|S_i|}{|S|} * Entropy(S_i)$$

Entropy(S_i)

Keterangan:

S = himpunan kasus

A = atribut

K = jumlah partisi atribut A

|S_i| = jumlah kasus pada partisi ke-i

|S| = jumlah kasus dalam S

Gain ratio

Untuk melakukan perhitungan *gain*
Ratio terlebih dahulu menghitung nilai
Split Information (Madadipouya, 2015).

Split Information =

$$- \sum_{i=1}^c \frac{S_i}{S} \log_2 \frac{S_i}{S}$$

Keterangan:

A merupakan pecahan hasil dari c
yang di sebit si sampai dengan sc yang
memiliki nilai banyak dengan menyatakan
rumus *gain ratio*.

$$Gain Rasio = \frac{Gain(S,A)}{Split Information(S,A)}$$

Gini index:

$$Gain(A) = 1 - \sum_{k=1}^c PK^2$$

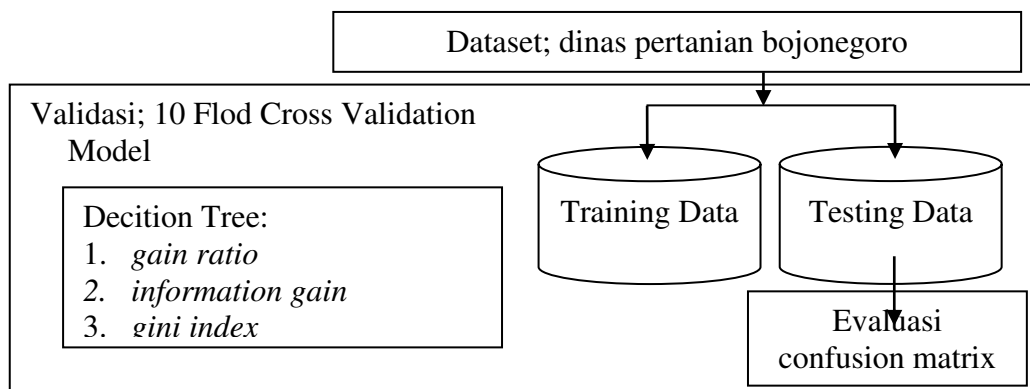
Keterangan :

K = atribut

C = Jumlah variabel Y

Pk = Proporsi jumlah kelas atribut K
terhadap jumlah kelas C.

Penerapan algoritma *Decision tree*
sebagai klasifikasi mutu produksi tanaman
jagung ditunjukkan pada Gambar 1.



Gambar 2. Model penelitian yang diusulkan

Tahapan metode dalam penelitian ini adalah:

1. Penerapan Metode *Decision tree*
2. Validasi menggunakan *k-fold cross validation*
3. Evaluasi menggunakan *Confusion Matrix*

Evaluasi dengan *confusion matrix*

Evaluasi menggunakan *Confusion Matrix* menghasilkan nilai akurasi, presisi, *recall*. Akurasi pada klasifikasi merupakan evaluasi hasil ketepatan data secara benar setelah dilakukan pengujian terhadap metode yang digunakan. *Precision* atau *confidence* merupakan jumlah kasus yang diprediksi negatif maupun positif. *Recall* atau *sensitivity* merupakan jumlah kasus positif yang sebenarnya diprediksi positif secara benar (Andriani, 2013; Berndtsson dkk., 2007).

Tabel 1. *Confusion matrix*

Actual	Predicted	
	-	+
-	True negative	False positive
+	False negative	True positive

positive merupakan jumlah *record* positif pada dataset klasifikasi maupun prediksi *true positive*. *negative* merupakan jumlah *record negative* pada dataset klasifikasi maupun prediksi *true negative*. *False positive* merupakan jumlah *record positive* pada dataset klasifikasi maupun

prediksi *positive*. *False negative* merupakan jumlah *record positive* pada dataset klasifikasi maupun prediksi *negative*.

Berikut adalah persamaan rumus *confusion matrix* :

1. Akurasi (Acc): jumlah prediksi yang benar, Persamaan di bawah adalah persamaan akurasi.

$$akurasi = \frac{tp+tn}{tp+tn+fp+fn}$$

2. *Sensitivity*: membandingkan jumlah tp terhadap tupel yang positif. Persamaan di bawah adalah persamaan *sensitivity*.

$$sensitivity = \frac{tp}{tp+fn}$$

3. *Specificity*: membandingkan jumlah tn terhadap tupel yang negative. Persamaan di bawah adalah persamaan *specificity*.

$$specificity = \frac{tn}{tn+fp}$$

4. *Positive predictive value*: jumlah kasus dengan hasil diagnosis positif. Persamaan di bawah adalah persamaan *positive predictive value*.

$$positive\ predictive\ value = \frac{tp}{tp+fp}$$

5. *Negative Predictive Value*: jumlah kasus dengan hasil diagnosis *negative*.

Persamaan di bawah adalah persamaan *negative predictive value*.

$$\text{positive predictive value} = \frac{tn}{tn+fn}$$

HASIL DAN PEMBAHASAN

Pengumpulan data

Data dalam penelitian ini adalah data tanaman jagung yang diperoleh dari database dinas pertanian Kabupaten Bojonegoro. Dataset ini terdiri dari 4961 data dengan 10 atribut, yaitu :

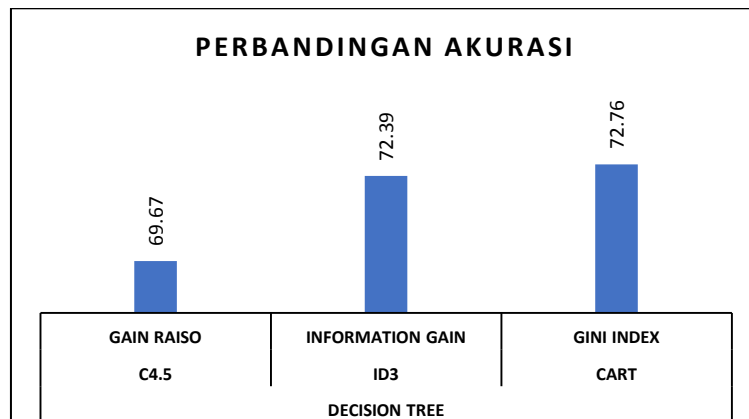
1. Varietas sebagai *regular attribute*.
2. Panjang sebagai *regular attribute*.
3. Bentuk sebagai *regular attribute*.
4. Warna sebagai *regular attribute*.
5. Rasa sebagai *regular attribute*.

6. Teknik sebagai *regular attribute*.
7. Musim sebagai *regular attribute*.
8. Hama sebagai *regular attribute*.
9. PH sebagai *regular attribute*.
10. Mutu sebagai *label*

Tabel 2. Perbandingan akurasi *decision tree*

C4.5	ID3	CART
<i>gain raiso</i>	<i>information gain</i>	<i>gini index</i>
69.67	72.39	72.76

Setelah dilakukan pengukuran terhadap kriteria tahap selanjutnya dilakukan proses membandingkan akurasi dari masing-masing kriteria yang digunakan.



Gambar 3. grafik perbandingan akurasi *decision tree*

Dari grafik perbandingan diatas penggunaan kriteria dari *gain ratio*, *information gain* , dan *gini index* pada *decision tree* menunjukkan perbedaan yang signifikan, *gini index* menunjukkan tingkat akurasi yang tertinggi, sedangkan akurasi terendah pada kriteria *gain ratio*.

KESIMPULAN

Hasil penelitian dari ketiga kriteria yang sudah diimplementasikan pada *decision tree* yaitu *gain ratio*, *information gain* dan *gini index*. Hasil akurasi menunjukkan bahwa nilai kriteria tertinggi

sebesar 72,76% pada kriteria *gini index* dan nilai terendah pada kriteria *gain ratio* sebesar 60,67%.

DAFTAR PUSTAKA

- Andriani, A. (2013). Sistem pendukung keputusan berbasis decision tree dalam pemberian beasiswa studi kasus: Amik BSI Yogyakarta. *Seminar Nasional Teknologi Informatika Dan Komunikasi, 2013*, 163–168.
- Berndtsson, M., Hansson, J., Olsson, B., & Lundell, B. (2007). *Thesis projects: A guide for students in computer science and information systems*. Springer Science & Business Media.
- Bustomi, M. A., & Dzulfikar, A. Z. (2014). Analisis distribusi intensitas RGB citra digital untuk klasifikasi kualitas biji jagung menggunakan jaringan syaraf tiruan. *Jurnal Fisika dan Aplikasinya*, 10(3), 127–132.
- Dawson, C. W. (2009). *Projects in computing and information systems: A student's guide*. Pearson Prentice Hall.
- Firmansyah, I. U., Aqil, M., & Sinuseng, Y. (2007). *Penanganan pascapanen jagung*. Puslitbang Tanaman Pangan, Badan Litbang Pertanian.
- Fitriani, D., Hasbullah, R., & Rachmat, R. (2017). Penentuan prioritas sarana pascapanen jagung untuk menurunkan kehilangan hasil dengan metode Analytical Hierarchy Process (AHP). *Jurnal Penelitian Pascapanen Pertanian*, 12(2), 60–69.
- Hamdi, A. S., & Bahrudin, E. (2015). *Metode penelitian kuantitatif aplikasi dalam pendidikan*. Deepublish.
- Kusumaningtyas, S., & Asmara, R. A. (2016). Identifikasi Kematangan Buah Tomat Berdasarkan Warna Menggunakan Metode Jaringan Syaraf Tiruan (JST). *Jurnal Informatika Polinema*, 2(2), 72–72.
- Madadipouya, K. (2015). A new decision tree method for data mining in medicine. *Advanced Computational Intelligence: An International Journal (ACII)*, 2(3).
- Prabowo, R. Y., Rahmadwati, R., & Mudjirahardjo, P. (2018). Klasifikasi kandungan nitrogen berdasarkan warna daun melalui color clustering menggunakan metode fuzzy c-means dan hybrid PSO k-means. *Jurnal EECCIS*, 12(1), 1–8.
- Prasetyo, E. (2014). *Data mining mengolah data menjadi informasi menggunakan matlab*. Andi Offset.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81–106. <https://doi.org/10.1023/A:1022643204877>
- Safuan, L. O., & Hadini, H. (2012). Klasifikasi genotip jagung lokal asal kabupaten wakatobi dan kabupaten bombana berdasarkan karakter fenotipnya. *J. Agroteknos*, 2(3), 126–133.
- Siregar, G., & Nugraha, S. (2018). Perkembangan produksi dan konsumsi jagung di provinsi Sumatera Utara. *Journal Agribusiness Sciences (JASC)*, 1(01). <http://jurnal.umsu.ac.id/index.php/JASC/article/view/1542>
- Yaumi, M. (2017). *Prinsip-prinsip desain pembelajaran: Disesuaikan dengan kurikulum 2013 edisi Kedua* (2 ed.). Kencana.