

BAB II

LANDASAN TEORI

2.1 Budidaya Ayam Broiler (Ras pedaging)

Ayam ras pedaging atau ayam broiler merupakan bangsa unggas yang arah kemampuan utamanya adalah untuk menghasilkan daging yang banyak dengan kecepatan pertumbuhan yang sangat pesat. Dalam waktu 5-6 minggu ayam broiler sudah memiliki bobot tubuh hingga 2 kg. Ayam ini merupakan jenis ras unggulan hasil persilangan dari bangsa-bangsa ayam yang memiliki produktivitas tinggi terutama dalam memproduksi daging [Rasyaf, M. 2003].

Dalam usaha ternak ayam broiler, peternak harus mengusahakan agar ternaknya tetap hidup dengan memenuhi segala kebutuhan hidup ternaknya. Makanan sebagai syarat utama harus diberikan agar kebutuhan nutrisi terpenuhi. Tanpa makanan yang sesuai dengan kebutuhan nutrisi untuk hidup dan produksi, alat produksi ini tidak akan bekerja baik. Bahkan kematian dapat terjadi akibat pemberian makanan yang tidak sesuai dengan kebutuhannya.

Lokasi untuk peternakan tidak berada di dalam kota atau di tepi kota. Lokasi ini harus memenuhi beberapa persyaratan antara lain :

1. Lokasi yang cukup jauh dari keramaian/perumahan penduduk.
2. Lokasi mudah terjangkau dari pusat-pusat pemasaran.
3. Lokasi terpilih bersifat menetap, artinya tidak mudah terganggu oleh keperluan-keperluan lain selain untuk usaha peternakan.

Keunggulan ayam broiler akan terbentuk bila didukung oleh lingkungan karena sifat genetik saja tidak menjamin keunggulan itu akan terlihat. Menurut (Rasyaf, M. 2004) hal-hal yang mendukung keunggulan ayam broiler adalah sebagai berikut :

1. Makanan sebaiknya memperhatikan kualitas dan kuantitas dalam pemberiannya. Pertumbuhan yang sangat cepat tidak akan tampak bila tidak didukung dengan ransum yang mengandung protein dan asam amino yang seimbang sesuai kebutuhan ayam.

2. Temperatur lingkungan Ayam broiler akan tumbuh optimal pada temperatur lingkungan 19°-21°c. Temperatur lingkungan di Indonesia lebih panas, apalagi di daerah pantai sehingga ayam akan mengurangi beban panas dengan banyak minum dan tidak makan. Akibatnya, sejumlah unsur nutrisi dan keperluan nutrisi utama yang berasal dari makanan menjadi tidak masuk ke dalam tubuh ayam. Jadi, temperatur ini secara tidak langsung berpengaruh terhadap kemampuan ayam broiler untuk bertahan hidup.
3. Pemeliharaan Bibit yang baik membutuhkan pemeliharaan yang baik pula. Ayam memerlukan perawatan dan makanan yang baik. Perawatan ini termasuk vaksinasi yang baik dan benar. Jika vaksinasinya tidak benar maka akan timbul penyakit yang akan mengakibatkan kematian.
4. Pemilihan DOC (Day Old Chicken) DOC adalah anak ayam umur sehari yang akan dibesarkan dan dipelihara menjadi ayam ras pedaging. Dalam memilih bibit DOC yang baik ada beberapa pedoman yang harus diperhatikan yakni:
 - a. Anak ayam (DOC) berasal dari induk yang sehat.
 - b. Bulu tampak halus dan penuh serta baik pertumbuhannya.
 - c. Tidak terdapat kecacatan pada tubuhnya.
 - d. Anak ayam mempunyai nafsu makan yang baik.
 - e. Ukuran badan normal, ukuran berat badan antara 35-40 gram.
 - f. Tidak ada letakan tinja diduburnya.

2.2. Definisi Sistem

Sistem secara fisik adalah kumpulan dari elemen-elemen yang beroperasi bersama-sama untuk menyelesaikan suatu sasaran (Gordon, 1991).

Menurut (Jagianto, 2005) Sistem adalah kumpulan dari elemen-elemen yang berinteraksi untuk mencapai suatu tujuan tertentu.

2.3. Data Mining

Secara sederhana data mining adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar. Data mining juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. Data mining, sering juga disebut sebagai knowledge discovery in database (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar

Data mining adalah kegiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam database, data warehouse, atau penyimpanan informasi lainnya. Data mining berkaitan dengan bidang ilmu – ilmu lain, seperti database system, data warehousing, statistik, machine learning, information retrieval, dan komputasi tingkat tinggi. Selain itu, data mining didukung oleh ilmu lain seperti neural network, pengenalan pola, spatial data analysis, image database, signal processing. Data mining didefinisikan sebagai proses menemukan pola-pola dalam data. Proses ini otomatis atau seringnya semiotomatis. Pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan, biasanya keuntungan secara ekonomi. Data yang dibutuhkan dalam jumlah besar (Huda, 2010).

2.3.1. Tahapan Metode Data Mining

Sebagai suatu rangkaian proses, data mining dapat dibagi menjadi beberapa tahap yang diilustrasikan di Gambar, Tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung dengan perantaraan *knowledge base*. Tahap – tahap data mining ada 6 yaitu :

1. Pembersihan data (*data Cleaning*)

Pembersihan data merupakan proses menghilangkan noise dan data yang tidak konsisten atau data yang tidak relevan. Pada umumnya data diperoleh, baik dari *database* suatu perusahaan maupun hasil eksperimen,

memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevan dengan hipotesa data mining yang dimiliki. Data-data yang tidak relevan itu juga lebih baik dibuang. Pembersihan data juga akan mempengaruhi performansi dari teknik data mining karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

2. Intregasi data (*data integraton*)

Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru. Tidak jarang data yang diperlukan untuk data mining tidak hanya berasal dari satu *database* atau file teks. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada.

3. Seleksi data (*data selection*)

Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*.

4. Transformasi data (*data transformation*)

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining. Beberapa metode data mining membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa metode standar seperti analisis asoisasi dan clustering hanya bisa menerima input data kategorikal, Karenanya data berupa angka numerik yang berlanjut perlu dibagi-bagi menjadi beberapa interval. Proses ini sering disebut transformasi data.

5. Proses mining

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

6. Evaluasi pola (*pattren evaluation*)

Untuk mengidentifikasi pola-pola menarik kedalam *knowledge based* yang ditemukan. Dalam tahap ini hasil dari teknik data mining berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki proses data mining, mencoba metode data mining yang lebih sesuai, atau menerima hasil ini sebagai suatu hal yang diluar dugaan yang mungkin bermanfaat. (Rizal, 2015)

2.3.2. Fungsi Data Mining

Data mining bukan hanya pelengkap saja dalam suatu database, melainkan mempunyai fungsi yang yang penting untuk membantu penggunanya mendapatkan informasi yang berguna serta meningkatkan pengetahuan bagi sang pengguna itu sendiri dan dapat nantinya berguna untuk orang banyak. Pada dasarnya, data mining mempunyai empat fungsi dasar yaitu: (Romario,2013)

1. Fungsi Klasifikasi (*classification*) Data mining dapat digunakan untuk mengelompokkan data-data yagn jumlahnya besar menjadi data-data yang lebih kecil.
2. Fungsi Segmentasi (*Segmentation*) Disini data mining juga digunakan untuk melakukan segementasi (pembagian) terhadap data berdasarkan karakteristik tertentu.
3. Fungsi Asosiasi (*Association*) Disini data mining juga digunakan untuk mencari hubungan antara karakteristik tertentu .
4. Fungsi pengurutan (*Sequencing*) Pada Fungsi ini, data mining digunakan untuk mengidentifikasikan perubahan pola yang telah terjadi dalam jangka waktu yang tertentu.

2.4. Klasifikasi

Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang tersedia. Dalam klasifikasi ada dua pekerjaan utama yang dilakukan, yaitu (1) pembangunan model sebagai prototipe untuk disimpan sebagai memori dan (2) penggunaan model tersebut untuk melakukan pengenalan/klasifikasi/prediksi pada suatu objek data lain agar diketahui dikelas mana objek data tersebut dalam model yang sudah disimpannya. Model dalam klasifikasi mempunyai arti yang sama dengan kotak hitam, dimana ada suatu model yang menerima masukan, kemudian mampu melakukan pemikiran terhadap masukan tersebut dan memberikan jawaban sebagai keluaran dari hasil pemikirannya. (Presetyo, E. 2012)

Tahapan dari klasifikasi dalam data mining menurut (Han dan Kamber, 2006) terdiri dari :

a. Pembangunan Model

Pada tahapan ini dibuat sebuah model untuk menyelesaikan masalah klasifikasi class atau atribut dalam data. Tahap ini merupakan fase pelatihan, dimana data latih dianalisis menggunakan algoritma klasifikasi, sehingga model pembelajaran direpresentasikan dalam bentuk aturan klasifikasi.

b. Penerapan Model

Pada tahapan ini model yang sudah dibangun sebelumnya digunakan untuk menentukan atribut/class dari sebuah data baru yang atribut/classnya belum diketahui sebelumnya. Tahap ini digunakan untuk memperkirakan keakuratan aturan klasifikasi terhadap data uji. Jika model dapat diterima, maka aturan dapat diterapkan terhadap klasifikasi data baru.

2.5. Pengertian Decision Tree

Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Selain itu dapat diekspresikan dalam

bentuk bahasa basis data seperti *Structure Query Language* untuk mencari *record* pada kategori tertentu. (Kusrini dan Emha, 2009).

Pohon keputusan juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan variabel target.

Sebuah pohon keputusan adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan *record* yang lebih kecil dengan menerapkan serangkaian aturan keputusan, dengan masing-masing rangkaian pembagian, anggota himpunan hasil menjadi mirip satu dengan yang lain. (Rizal, 2015)

2.5.1. Jenis – Jenis Decision Tree

Beberapa model *decision tree* yang sudah dikembangkan antara lain C4.5 atau ID3 dan CART. Berikut ini akan dijelaskan model dari decision tree tersebut :

1. C4.5 atau ID3

Decision Tree menggunakan algoritma ID3 atau C4.5, yang diperkenalkan dan dikembangkan pertama kali oleh Quinlan yang merupakan singkatan dari *Iterative Dichotomiser 3* atau *Induction of Decision 3*. Algoritma ID3 membentuk pohon keputusan dengan metode *divide-and-conquer* data secara rekursif dari atas ke bawah. Strategi pembentukan Decision Tree dengan algoritma ID3 adalah:

- A. Pohon dimulai sebagai *node* tunggal (akar/*root*) yang merepresentasikan semua data.
- B. Sesudah *node root* dibentuk, maka data pada *node* akar akan diukur dengan *information gain* untuk dipilih atribut mana yang akan dijadikan atribut pembaginya.
- C. Sebuah cabang dibentuk dari atribut yang dipilih menjadi pembagi dan data akan didistribusikan ke dalam cabang masing-masing.
- D. Algoritma ini akan terus menggunakan proses yang sama atau bersifat rekursif untuk dapat membentuk sebuah *Decision Tree*. Ketika sebuah atribut telah dipilih menjadi *node* pembagi atau cabang, maka atribut

tersebut tidak diikuti lagi dalam penghitungan nilai *information gain*.

- E. Proses pembagian rekursif akan berhenti jika salah satu dari kondisi dibawah ini terpenuhi :
- a. Semua data dari anak cabang telah termasuk dalam kelas yang sama.
 - b. Semua atribut telah dipakai, tetapi masih tersisa data dalam kelas yang berbeda. Dalam kasus ini, diambil data yang mewakili kelas yang terbanyak untuk menjadi label kelas pada *node* daun. Tidak terdapat data pada anak cabang yang baru. Dalam kasus ini, *node* daun akan dipilih pada cabang sebelumnya dan diambil data yang mewakili kelas terbanyak untuk dijadikan label kelas.

Metode C4.5 dan ID3 memiliki perbedaan dalam nilai tiap atribut. Metode C4.5 menggunakan atribut yang bernilai kategorikal dan numerikal, sedangkan metode ID3 menggunakan atribut yang bernilai kategorikal. Metode *decision tree C4.5* inilah yang digunakan dalam tugas akhir ini.

2. CART

CART adalah singkatan dari *Classification And Regression Tree*. Dalam CART ada dua langkah penting yang harus diikuti untuk mendapatkan *tree* dengan performansi yang optimal. Yang pertama adalah pemecahan objek secara berulang berdasarkan atribut tertentu. Yang kedua, *prunning* (pemangkasan) dengan menggunakan data validasi.

Misalkan kita mempunyai variabel independent $x_1, x_2, x_3, \dots, x_n$ dan variabel dependent atau output y . Pemecahan secara berulang berarti kita bagi objek ke dalam kotak-kotak berdasarkan nilai variabel x_1, x_2 atau x_r . Cara ini diulang sehingga dalam suatu kotak sebisa mungkin berisi observasi dalam kelompok atau kelas yang sama.

Langkah berikutnya sesudah dilakukan pemecahan objek atau data secara berulang adalah melakukan *prunning*. Dalam *prunning* kita ingin memangkas *tree* yang mungkin terlalu besar dan terjadi fenomena

overfitting. *Overfitting* merupakan sebuah satu buah pengelompokan yang mungkin hanya berisi satu data yang memungkinkan data tersebut merupakan *noise* yang ada di data training dan bukan pola yang mungkin terjadi dalam data testing atau data validasi. *Prunning* terdiri dari beberapa langkah pemilihan secara berulang simpul yang akan dijadikan simpul daun. Dengan mengubah simpul menjadi simpul daun artinya tidak akan dilakukan pemecahan lagi sesudah itu. Dengan demikian ukuran *tree* akan berkurang (Sari, 2017)

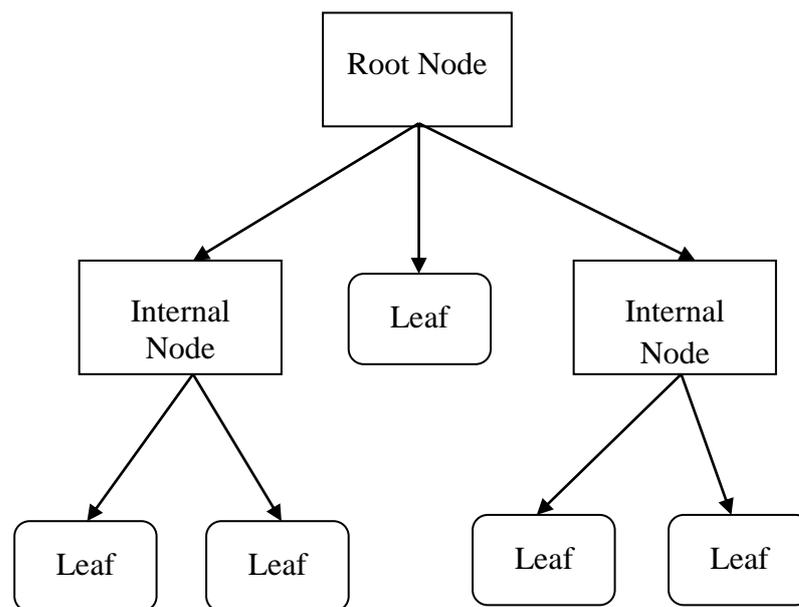
2.5.2. Model Decision Tree

Decision tree adalah *flow-chart* seperti *struktur tree*, dimana tiap *internal node* menunjukkan sebuah test pada sebuah atribut, tiap cabang menunjukkan hasil dari test, dan *leaf node* menunjukkan *class-class* atau *class distribution*.

Selain karena pembangunannya relatif cepat, hasil dari model yang dibangun mudah untuk dipahami. Pada *decision tree* terdapat 3 jenis *node*, yaitu:

- a. *Root Node*, merupakan *node* paling atas, pada *node* ini tidak ada *input* dan bisa tidak mempunyai *output* atau mempunyai *output* lebih dari satu.
- b. *Internal Node*, merupakan *node* percabangan, pada *node* ini hanya terdapat satu *input* dan mempunyai *output* minimal dua.
- c. *Leaf node* atau *terminal node*, merupakan *node* akhir, pada *node* ini hanya terdapat satu *input* dan tidak mempunyai *output*. (Rizal,2015)

Contoh dari model pohon keputusan yaitu seperti pada **gambar 2.1** berikut:



Gambar 2.1 Model *Decision Tree*

2.6. Algoritma Decision Tree C4.5

Algoritma C4.5 diperkenalkan oleh Quinlan pada tahun 1996 sebagai versi perbaikan dari ID3. Dalam ID3, induksi decision tree hanya bisa dilakukan pada fitur bertipe kategorikal (nominal atau ordinal), sedangkan tipe numerik (interval atau rasio) tidak dapat digunakan.

Yang menjadi hal penting dalam induksi decision tree adalah bagaimana menyatakan syarat pengujian pada node. Ada 3 kelompok penting dalam syarat pengujian node :

1. Fitur biner

Adalah Fitur yang hanya mempunyai dua nilai berbeda. Syarat pengujian ketika fitur ini menjadi node (akar maupun interval) hanya punya dua pilihan cabang.

2. Fitur kategorikal

Untuk fitur yang nilainya bertipe kategorikal (nominal atau ordinal) bisa mempunyai beberapa nilai berbeda. Secara umum ada 2 pemecahan yaitu pemecahan biner (*binary splitting*) dan (*multi splitting*).

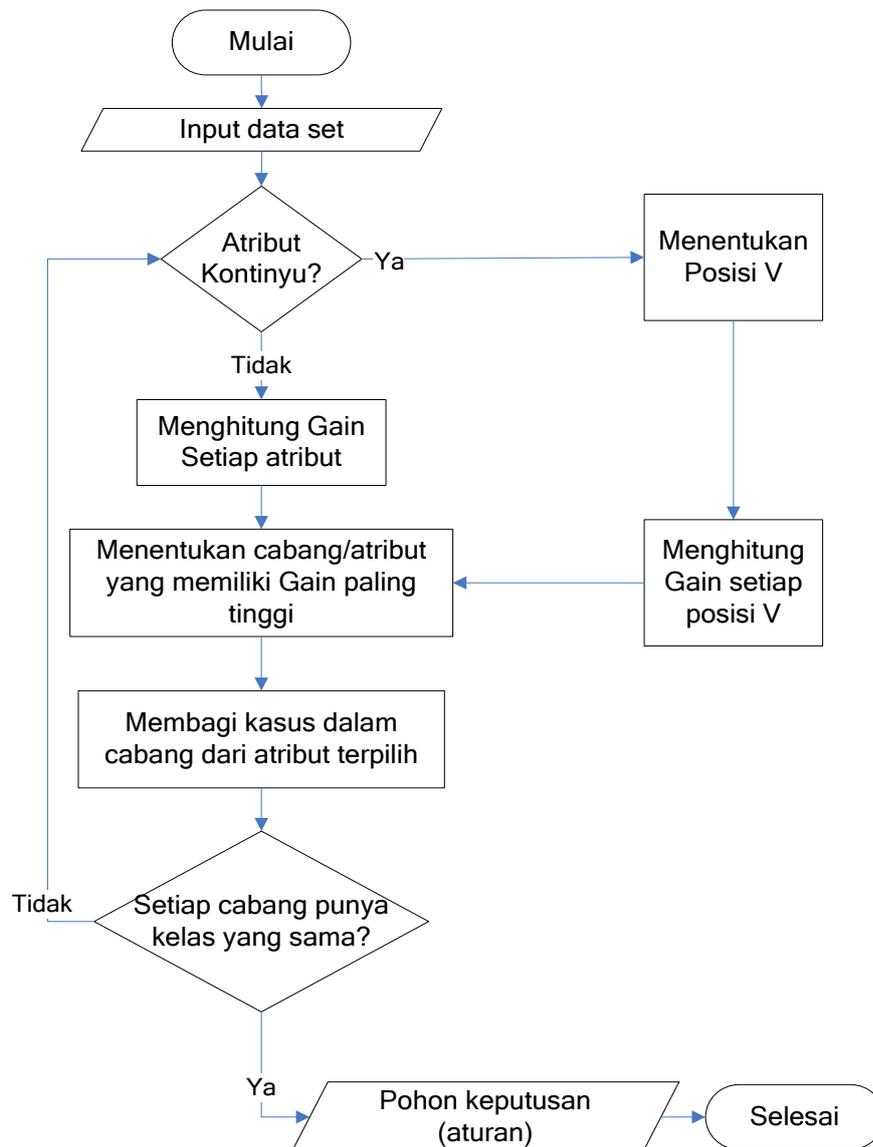
3. Fitur numerik

Untuk fitur bertipe numerik, Syarat pengujian dalam node (akar maupun internal) dinyatakan dengan pengujian perbandingan ($A \leq V$) atau ($A > V$) dengan hasil biner.

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Pilih atribut sebagai akar.
2. Buat cabang untuk tiap-tiap nilai.
3. Bagi kasus dalam cabang.
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Berikut ini akan dijelaskan secara lebih detail algoritma C4.5 menggunakan *flowchart* yang disajikan pada **gambar 2.2**.



Gambar 2.2 Flowchart algoritma Decision Tree C4.5

Untuk memilih atribut sebagai simpul akar (*root node*) atau simpul dalam (*internal node*), didasarkan pada nilai *information gain* tertinggi dari atribut-atribut yang ada. Sebelum perhitungan *information gain*, akan dilakukan perhitungan *entropy*. *Entropy* merupakan distribusi probabilitas dalam teori informasi dan diadopsi kedalam algoritma C4.5 untuk mengukur tingkat homogenitas distribusi kelas dari sebuah himpunan data (*data set*). Semakin tinggi tingkat *entropy* dari sebuah data maka semakin homogen distribusi kelas pada

data tersebut. Perhitungan *information gain* menggunakan rumus 2.1, sedangkan *entropy* menggunakan rumus 2.2.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \dots\dots(2.1)$$

dimana,

S : Himpunan kasus

A : Atribut

n : Jumlah partisi atribut A

|S_i|: Jumlah kasus pada partisi ke i

|S| : Jumlah kasus dalam S

$$Entropy(S) = - \sum_{i=1}^n p_i * \log_2 p_i \dots\dots\dots(2.2)$$

dimana,

S : Himpunan kasus

n : Jumlah partisi S

p_i : Proporsi dari S_i terhadap S

Selain *Information Gain* kriteria yang lain untuk memilih atribut sebagai pemecah adalah *Rasio Gain*. Perhitungan rasio gain menggunakan rumus 2.3, sedangkan split information menggunakan rumus 2.4.

$$GainRasio(S, A) = \frac{Gain(S, A)}{SplitInformation(S, A)} \dots\dots\dots(2.3)$$

$$SplitInformation(S, A) = - \sum_{i=1}^c \frac{S_i}{S} \log_2 \frac{S_i}{S} \dots\dots\dots(2.4)$$

dimana S₁ sampai S_c adalah c subset yang dihasilkan dari pemecahan S dengan menggunakan atribut A yang mempunyai sebanyak c nilai. (Sari, 2017)

2.7. Penelitian Sebelumnya

Penelitian sebelumnya yang menggunakan metode Naïve Bayes dilakukan oleh di lakukan oleh Moh Wahyul Hadi Saputro, lulusan Tahun 2015 universitas muhammadiyah gresik teknik informatika. Penelitian untuk “**SISTEM** *Prediksi Untung Rugi Pada Hasil Peternakan Ayam Broiler Dengan Metode Klasifikasi Naive Bayes*”. Variabel yang digunakan meliputi keuntungan, modal sendiri, total modal, rentabilitas ekonomi, rentabilitas modal sendiri, strata 1, strata2, strata3. Tingkat keberhasilan plasma tiap periode produksi pada tingkat strata 2 adalah yang paling tinggi dibanding dengan strata 3 dan strata 1. Atribut yang terdapat pada tabel mewakili fitur data yang digunakan meliputi Musim, Populasi, Bibit, Obat, dan Pakan. Jumlah data yang digunakan sebanyak 572 *record* dengan kelas Baik dan Kurang. Penelitian ini diuji sebanyak 3 kali pengujian dengan rata-rata akurasi sebesar 71,24%.

Penelitian selanjutnya adalah tentang metode Decision Tree C4.5 dalam penelitian berjudul “Aplikasi Klasifikasi Penentuan Penerimaan Beras Miskin (Raskin) ini Ds.Sidomulyo Kec.Deket Kab.Lamongan dengan Metode Decision Tree C4.5”, dibuat oleh M.Basroni Rizal (Universitas Muhammadiyah Gresik, 2016). Penelitian dilakukan untuk mengklasifikasi penentuan penerimaan beras miskin. Dari 600 data kepala keluarga di Ds. Sidomulyo Kec. Deket Kab. Lamongan, data tersebut diambil 40 % yang akan dijadikan sebagai data uji dan 60 % akan menjadi data latih. Jadi jumlah pembagiannya adalah 240 data sebagai data latih dan 360 data untuk data uji. Penelitian ini diuji sebanyak 7 kali pengujian dengan rata-rata akurasi sebesar 92%.

Penelitian selanjutnya adalah tentang metode *Decision Tree C4.5* dalam penelitian yang berjudul “*System prediksi prestasi akademik mahasiswa menggunakan metode decision tree C4.5 (Studi kasus:Jurusan Teknik informatika UNMUH GRESIK)*”, dibuat oleh Aunur Rasyid (Universitas Muhammadiyah Gresik, 2014). Tujuan dari penelitian tersebut adalah untuk menghasilkan informasi perkiraan kategori prestasi mahasiswa menggunakan metode *Decision*

Tree C4.5 sebagai peringatan dini dan motivasi mahasiswa dalam mendapatkan prestasi yang maksimal. Atribut-atribut yang digunakan adalah instansi sekolah asal (SMK,SMA atau MA), satatus sekolah asal (Negri atau Swasta), jurusan sekolah asal (IPA,IPS,Bahasa,Teknik,Administrasi), nilai rata-rata UN, status kerja (Sudah atau Belum), dan pihak yang mempengaruhi mahasiswa dalam memilih kuliah (Sendiri,Orang tua atau Orang lain). Hasil dari penelitian tersebut, Sistim Prediksi mahasiswa yang dirancang menggunakan algoritma C4.5 dapat memprediksi prestasi mahasiswa agar mampu mempertahankan kondisinya atau melakukan perbaikan utuk mencapai prestasi yang maksimal. Hasil akurasi dari penelitian tersebut adalah 90%.