

## KLASIFIKASI HASIL PREDIKSI PANEN PADI BERDASARKAN FISILOGIS MENGGUNAKAN METODE *NAÏVE BAYES CLASSIFICATION*

Rudi Hariyanto<sup>1)</sup>, Anang Aris Widodo<sup>2)</sup>

<sup>1)</sup>Informatika, Universitas Merdeka Pasuruan, Pasuruan  
Email: [rudihariy4nt0@gmail.com](mailto:rudihariy4nt0@gmail.com)

<sup>2)</sup>Informatika, Universitas Merdeka Pasuruan, Pasuruan  
Email: [anangariswidodo@gmail.com](mailto:anangariswidodo@gmail.com)

### ABSTRAK

Dalam prediksi hasil panen padi, Perlu keahlian khusus dalam menganalisa hasil prediksi padi menjelang masa panen dengan mempertimbangkan banyak faktor seperti: luas lahan, faktor cuaca, kondisi air, kondisi tanaman padi dan lain-lain. Akan tetapi tidak semua para petani mengetahui perhitungan dalam memprediksi hasil panen padi. Karena keterbatasan ilmu pengetahuan, sehingga para petani tidak bisa menafsir hasilnya bahkan memberikan harga padi yang pantas disaat-saat jelang panen. Ahli pertanian dalam hal ini mempunyai kemampuan dalam prediksi hasil panen padi saat menjelang masa panen, sains dan teknologi ini yang nantinya dijadikan sebuah sistem dalam memberikan informasi hasil prediksi panen padi. Metode inference suatu mekanisme berfikir dan pola-pola penalaran yang digunakan oleh sistem untuk mencapai suatu kesimpulan. Metode ini akan menganalisa masalah tertentu dan selanjutnya akan mencari jawaban atau kesimpulan yang terbaik. Pada penelitian ini di terapkan metode Naive Bayes Classifier untuk memprediksi hasil panen dan mengelompokkan hasilnya, baik atau buruk.

**Kata kunci:** Klasifikasi, Hasil Panen, Pendukung Keputusan, Naive Bayes

### ABSTRACT

*In the prediction of rice yields, special expertise is needed in analyzing the prediction results of rice before the harvest period by considering many factors such as: land area, weather factors, water conditions, condition of rice plants and others. However, not all farmers know the calculations in predicting rice yields. Because of the limitations of science, farmers cannot interpret the results and even provide a reasonable price of rice before harvest time. Agricultural experts in this case have the ability to predict rice yields nearing the harvest, science and technology which will be used as a system to provide information on the prediction of rice yields. The inference method makes the thinking mechanism and the reasoning patterns used by the system to reach a conclusion. This method will analyze a particular problem and then look for the answer or the best conclusion. In this research the Naive Bayes Classifier method is applied to predict yields and classify the results, good or bad.*

**Keywords:** Classification, Harvest Results, Decision Support, Naive Bayes

## PENDAHULUAN

Dalam prediksi hasil panen padi, Perlu keahlian khusus dalam menganalisa hasil prediksi padi menjelang masa panen dengan mempertimbangkan banyak faktor seperti: luas lahan, faktor cuaca, kondisi air, kondisi tanaman padi dan lain-lain, akan tetapi tidak semua para petani mengetahui perhitungan dalam memprediksi hasil panen padi. Dikarenakan keterbatasan ilmu pengetahuan, sehingga para petani tidak bisa menafsir hasilnya bahkan memberikan harga padi yang pantas disaat-saat menjelang panen. Ahli pertanian dalam hal ini mempunyai kemampuan dalam prediksi hasil panen padi saat menjelang masa panen, sains dan teknologi ini yang nantinya dijadikan sebuah sistem dalam memberikan informasi hasil prediksi panen padi. Metode inference adalah mekanisme berfikir dan pola-pola penalaran yang digunakan oleh sistem untuk mencapai suatu kesimpulan. Metode ini akan menganalisa masalah tertentu dan selanjutnya akan mencari jawaban atau kesimpulan yang terbaik. Penalaran dimulai dengan mencocokkan kaidah-kaidah dalam basis pengetahuan dengan fakta-fakta yang ada dalam basis data (Turban, 2007).

Fokus pada penelitian ini penulis memprediksi hasil panen padi antara 1 minggu hingga 1 bulan sebelum masa panen, diharapkan petani bisa memprediksi apakah hasilnya bagus atau tidak dengan memanfaatkan teknologi komputer sebagai alat bantu manusia yang sifatnya memberatkan (Sutojo, 2013), sehingga pekerjaan yang dahulu masih didominasi oleh tenaga manusia, sekarang bisa dialihkan dengan tenaga mesin (komputer) (Kusrini, 2007).

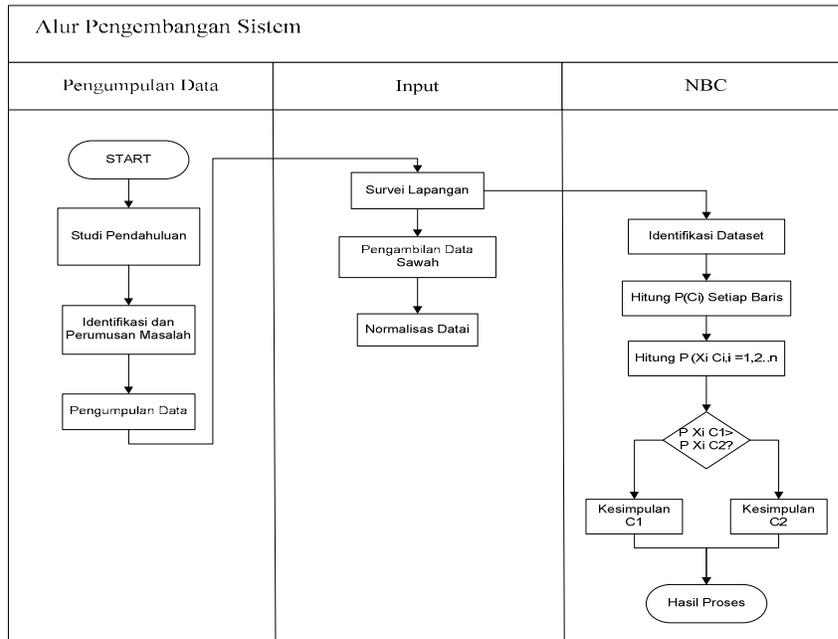
Mayoritas penelitian tersebut menggunakan metode klasifikasi dalam pengambilan kesimpulan seperti yang dilakukan seperti Setiawan dan Ratnasari 2014 mengklasifikasikan penyakit mata yang terdiri dari 52 gejala dan 15 jenis penyakit mata menggunakan metode *Naïve Bayes Classification* (NBC) dengan jumlah data uji coba sebanyak 12, dari hasil percobaan data tersebut diperoleh prosentase keberhasilan sebesar 83%. Kusumadewi, 2009 menggunakan algoritma *Naive Bayesian Classifier* untuk mengklasifikasikan status gisi pada seseorang menunjukkan kinerja sistem sebesar 93,2% dari 36 data. Anandita, 2015, Mengklasifikasi Tebu Dengan Menggunakan Algoritma *Naïve Bayes Classification* Pada Dinas Kehutanan Dan Perkebunan Pati kinerja sistem untuk masalah klasifikasi tebu produktif sebesar 73,3% dari 28 data.

Kecerdasan buatan Artificial Intelligence memberikan gambaran hasil prediksi padi agar lebih tepat dalam memberikan suatu kesimpulan dengan pemanfaatan teknologi komputer sebagai alat bantu secara tepat dan akurat (Sutojo, 2013). Dengan kemajuan teknologi yang semakin berkembang penelitian ini berinisiatif membangun Sistem Pendukung Keputusan dengan tujuan untuk memudahkan para petani dalam mengklasifikasikan memprediksi hasil panen padi sebelum masa panen.

## METODE PENELITIAN

Dalam melaksanakan penelitian yang dilakukan oleh penulis yaitu : tempat dan waktu penelitian, jenis penelitian. Adapun jenis penelitian adalah jenis penelitian kuantitatif (Linda, 2004). Dimana kuantitatif merupakan salah satu jenis penelitian yang spesifikasinya adalah sistematis, terencana, dan terstruktur dengan jelas sejak awal hingga pembuatan desain penelitiannya (Kusumadewi, S, 2003)

### Alur Pengembangan sistem NBC (*Naive Bayes Clasification*)



Gambar 1. Alur Sistem

Berdasarkan flowchart pada Gambar 1 dapat dijelaskan ke dalam beberapa tahap sebagai berikut :

1. *Studi Pendahuluan*

Studi pendahuluan sebagai tahap awal dalam metode penulisan ini. Di tahap ini penulis melakukan studi lapangan dengan terjun langsung pesawahan dan tengkulak padi (Pakar Tani) wilayah Pasuruan. Pengamatan langsung maupun wawancara dari beberapa tani atau ahlinya di bidang pertanian, adalah hal-hal yang dilakukan dengan tujuan mengetahui informasi-informasi awal mengenai kriteria kriteria tentang penyakit atau gejala dalam menentukan hasil panen.

2. *Identifikasi dan Perumusan Masalah*

Setelah melakukan penulisan pendahuluan, permasalahan yang dialami oleh setiap sawah dapat diidentifikasi melalui gejala-gejala perubahan fisiologis serta factor lain yang dialaminya, kemudian diklasifikasikan sesuai jenis baik dan buruk hasil panennya. Kemudian dip roses dengan metode NBC untuk membandingkan hasil kinerja serta tingkat keakurasiannya dalam memproses data.

3. *Pengumpulan Data*

Pada penelitian ini adalah data perubahan cuaca, faktor tanaman serta penyakit yang akan mengurahi hasil produktisi. Data-data tersebut berisi data dasar fisik tumbuhan, aliran acir dan faktor alam. Data didapatkan dari survei serta konsultasi dari para petani tau tengkulak di Pasuruan dalam rentan waktu Mei 2018 - Desember 2018.

4. *Pengambilan Data*

Dalam tahap ini data dari dokumen-dokumen survei serta wawancara dai beberapa petani maupun tengkulak padi di wilayah pasuruan lalu direkapitulasi kedalam aplikasi Microsoft Excel dengan bentuk kolom dan baris, kolom berisi data per fitur sedangkan baris merupakan data per sawah dan data akan disimpan dalam format .csv.

#### 5. *Normalisasi Data*

Normalisasi data ini bertujuan untuk menyeragamkan data untuk semua fitur. Data yang tidak berupa angka perlu dilakukan konversi menjadi data numerik. Misal data cuaca, hujan "1" untuk Hujan dan "0" untuk tidak Hujan, data jenis pohon yang umumnya berupa data tekstual seperti pohon roboh dan berdiri akan dikonversi menjadi "roboh=0" dan "berdiri=1". Selanjutnya semua data fitur akan dinormalisasi dalam rentang yang sama yaitu bernilai antara [0,1]. Sedangkan data pada kolom diagnosa hanya dirubah kedalam data angka saja karena ini menggambarkan kelas-kelas luaran, untuk Baik = "2", Sedang ="1" dan buruk = "0" (Arifianto, 2014).

#### 6. *Dataset*

Dataset adalah data yang telah dinormalisasi dan dimasukkan kedalam tabel basis data. Keseluruhan data akan dibagi menjadi dua yaitu:

##### a. Pembelajaran

Datasets yang digunakan sebagai data training dalam proses klasifikasi. Untuk data pembelajaran akan digunakan 90% dari 100 data keseluruhan.

##### b. Ujicoba

Dataset yang digunakan sebagai data pengujian dalam proses klasifikasi. Data ini harus benar-benar berbeda dari data pembelajaran untuk memenuhi aspek obyektifitas. Untuk data ujicoba akan digunakan 10% dari 100 data keseluruhan.

### **Tahapan Pembuatan Sistem**

Ada 7 tahapan pembuatan sistem yaitu sebagai berikut :

1. Perencanaan
2. Analisa
3. Desain

Setelah proses analisa selesai, selanjutnya adalah membuat desain (design).

Adapun macam – macam desain:

- a. *Flowchart*
- b. DFD (Data Flow Diagram)
- c. ERD (Entity Relationship Diagram)

4. Pengembangan
5. Testing
6. Implementasi
7. Pengoperasian dan Pemeliharaan

### **Variabel Penelitian**

Sesuai dengan paparan Sugiono, variabel penelitian adalah segala sesuatu yang ditetapkan peneliti untuk dipelajari sehingga diperoleh informasi tentang hal tersebut kemudian ditarik kesimpulan. Sesuai dengan penjelasan didalam latar belakang variabel dalam penelitian ini selanjutnya disebut fitur (Sugiono. 2009).

Tabel 1 Variabel/fitur penelitian

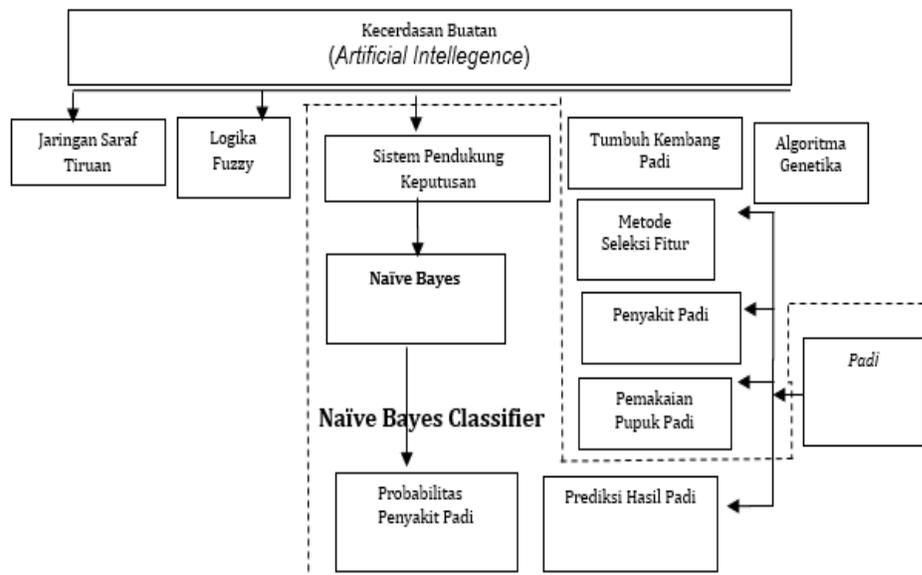
Dataset Penelitian	
Lahan	Panjang
	Lebar
	Luas
Isi Buah Padi	Buah di Pinggir
	buah di Tengah

Musim	Hujan
	Sedang
	Panas
Air	Kering
	Sedang
	Banjir
Daun	Kuning
	Agak coklat
	Coklat
Batang	Berdiri
	Sebagian Roboh
	Roboh

### HASIL DAN PEMBAHASAN

Sebuah metode dari data mining yang digunakan untuk memprediksi kategori atau kelas dari suatu data instance berdasarkan sekumpulan kriteria-kriteria dari data tersebut untuk memasukkan kedalam kelas tertentu dari sejumlah kelas yang tersedia (Daihani, DU., 2001). Pembangunan model prototipe untuk disimpan kedalam memori untuk melakukan klasifikasi/prediksi pada suatu objek agar dapat diketahui dikelas mana objek data tersebut disimpan. Input dari user diteruskan ke mesin inferensi untuk dilakukan pemrosesan lebih lanjut. Mesin inferensi sendiri merupakan kumpulan metodologi yang digunakan untuk melakukan penalaran terhadap informasi didalam basis data. Dengan penalaran tersebut diharapkan pemakai mendapatkan solusi yang sesuai dengan masalahnya. Inferensi merupakan proses menghasilkan kesimpulan berdasarkan fakta atau pengetahuan yang diketahui atau diasumsikan. Terdapat pendekatan untuk mengontrol inferensi dalam sistem pakar berbasis aturan pelacakan berdasarkan probabilitas (Naive Bayes).

### Konsep Solusi



Gambar 2. Skema seleksi untuk klasifikasi Hasil Panen Padi

### Naïve Bayes Classifier

Naïve Bayes Classification untuk klasifikasi prediksi hasil padi

Pada NBC setiap fitur direpresentasikan dalam pasangan atribut  $\langle a_1, a_2, \dots, a_n \rangle$  dimana  $a_1$  adalah fitur pertama,  $a_2$  adalah fitur kedua dan seterusnya. Sedangkan  $V$  adalah himpunan kategori kelas (hasil panen padi yang bagus). Pada saat klasifikasi, pendekatan Bayes akan menghasilkan label kategori yang paling tinggi probabilitasnya ( $V_{MAP}$ ) dengan masukan atribut  $\langle a_1, a_2, \dots, a_n \rangle$

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

Keterangan :

$X$  = Data dengan class yang belum diketahui

$Y$  = baik data  $X$  merupakan suatu class spesifik

$P(Y|X)$  = Probabilitas baik  $Y$  berdasar kondisi  $X$  (posteriori probability)

$P(Y)$  = Probabilitas baik  $Y$  (prior probability)

$P(X|Y)$  = Probabilitas  $X$  berdasarkan kondisi pada fitur

$$V_{MAP} = \arg v_{j \in V} \max P(Y_j | a_1, a_2, \dots, a_n) \quad (1.1)$$

Teorema Bayes menyatakan

$$V_{MAP} = \frac{P(A|B)P(B)}{P(A)}$$

Menggunakan teorema Bayes ini, persamaan (2.23) ini dapat ditulis :

$$V_{MAP} = \arg v_{j \in V} \max \frac{P(a_1, a_2, \dots, a_n | V_j)P(V_j)}{P(a_1, a_2, \dots, a_n)} \quad (1.2)$$

$P(a_1, a_2, \dots, a_n)$  nilai konstan untuk semua  $V_j$  sehingga persamaan ini dapat ditulis sebagai berikut:

$V_{MAP} = \arg v_{j \in V} \max p(a_1, a_2, \dots, a_n | V_j)$  menjadi tinggi karena jumlah fitur  $P(a_1, a_2, \dots, a_n | v_j)$  bisa jadi akan sangat besar. Ini disebabkan jumlah fitur tersebut sama dengan jumlah kombinasi posisi fitur dikali dengan jumlah kelas kategori. Naïve Bayes Classifier menyederhanakan hal ini dengan mengasumsikan bahwa dalam setiap kategori, setiap kata independen satu sama lain. Dengan kata lain :

$$P(a_1, a_2, \dots, a_n) = \prod_i P(a_i | v_j) \quad (1.3)$$

Substitusi persamaan ini dengan persamaan 1.2 akan menghasilkan :

$$V_{MAP} = \arg v_{j \in V} \max P(V_j) \prod_i P(a_i | V_j)$$

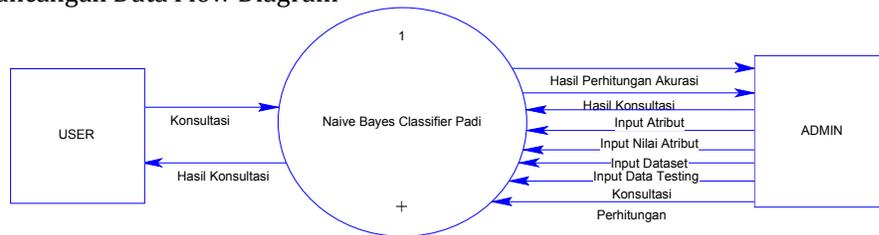
$P(v_j)$  dan probabilitas setiap fitur untuk setiap kategori  $P(w_k | v_j)$  dihitung pada saat pelatihan.

$$P(V_j) = \frac{|docs_j|}{|Contoh|} \quad (1.4)$$

$$P(V_j) = \frac{nk+1}{n+1Baikl}$$

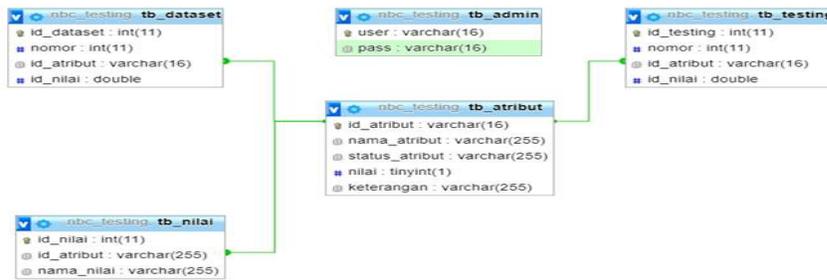
Di mana  $|docs_j|$  adalah jumlah fitur pada kategori  $j$  dan yang digunakan dalam pelatihan. Sedangkan  $nk$  adalah jumlah kemunculan fitur  $w_k$  pada kategori  $v_j$  dan  $|kelas|$  pada semua data latihan. Jumlah setiap fitur dalam tiap kelas dinyatakan sebagai  $n$ . Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas  $Y$  (Posterior) peluang munculnya kelas  $Y$  (sebelum masuknya sampel tersebut, seringkali disebut prior), dikali dengan peluang kemunculan karakteristik sampel pada kelas  $Y$  (disebut juga likelihood), dibagi dengan peluang kemunculan karakteristik sampel secara global (disebut juga evidence)

### Perancangan Data Flow Diagram



Gambar 3. Alur Sistem dalam DFD

### Relasi Antar Tabel



Gambar 4. Relasi Antar Tabel dalam ERD

Terkait dengan penelitian ini. Pembelajaran untuk dataset yang diseleksi telah selesai serta mendapatkan bobot-bobot terbaik, maka untuk mengetahui keberhasilan dan performa dari dua algoritma sistem harus dilakukan pengujian. Pengujian meliputi tingkat *akurasi*, *sensitifitas* dan *waktu*. Tahap ini menggunakan dataset ujicoba. Pengujian ini dilakukan untuk mengetahui pengaruh seleksi fitur data prediksi padi dalam proses klasifikasi

Probabilitas	A07	A08	A09	A10
Sedang	0.0065	0.0103	0.0016	0
Jelek	0.0061	0.01	0.0016	0

Kelompok	A07		A08		A09		A10	
	Hujan	Kemarau	Panas	Kering	Sedang	Banjir	Kuning	Agak coklat
Bagus	0.15384615384615	0	0.84615384615385	1	0	0	1	0
Sedang	0	0.79487179487179	0.20512820512821	0	1	0	0	1
Jelek	0.80952380952381	0.19047619047619	0	0	0	1	0	0

Kelompok	A01	A02	A03	A04	A05	A06	A11	A07 (Panas)	A08 (Kering)	A09 (Kuning)
	Bagus (0.325)	0.007617442052418	0.0095060318807285	0.0015094660323445	1.2288283373073	0.64332661443116	1.1545688711005	0.00042152200470364	0.84615384615385	1
Sedang (0.325)	0.0064959803960219	0.010250995292471	0.001623655582041	0	0.80756682789016	1.6151326757803	0.0005652082109352	0.20512820512821	0	0
Jelek (0.35)	0.006080859331423	0.009977741091204	0.0016175880076878	0	0.79452728557912	1.5890545711582	0.0017819071672886	0	0	0

Berdasarkan perhitungan, dengan Panjang Lahan: 20, Lebar Lahan: 30, Luas Lahan: 600, Isi Buah Pinggir: 12, Isi Buah Tengah: 10, Isi Buah Rata-Rata: 11, Musim: Panas, Kondisi Air: Kering, Kondisi Daun: Kuning, Kondisi Batang: Berdiri, Perhitungan: nbc, maka hasilnya: Bagus.

Gambar 5. Form Interface Hasil Pembobotan

## KESIMPULAN

Berdasarkan penjelasan yang terdapat pada bab sebelumnya maka pada penelitian ini dapat ditarik kesimpulan sebagai berikut :

1. Dengan adanya sistem ini, dapat membantu user khususnya petani padi dalam melakukan analisa yang lebih cepat secara komputerisasi dalam menghasilkan presentase hasil prediksi padi berdasarkan kriteria padi menjelang masa panen.
2. Presentase hasil analisa hasil prediksi panen padi dengan menggunakan proses perhitungan metode Naïve Bayes Clasivikasi sangat dipengaruhi hasilnya berdasarkan hasil pengujian, disimpulkan bahwa dari 10 orang total responden sebesar 90 persen. Hal ini menunjukkan bahwa penilaian terhadap sistem oleh responden yang paham terhadap hasil perhitungan padi.

## UCAPAN TERIMA KASIH

Peneliti menyampaikan ucapan terima kasih kepada pihak-pihak yang telah membantu pelaksanaan penelitian hingga selesai. Di antaranya Fakultas Teknologi Informasi Universitas Merdeka Pasuruan, LPPM Universitas Merdeka Pasuruan, Dinas Pertanian Kota dan Kabupaten Pasuruan.

## REFERENSI

- Arifianto. 2014. "Klasifikasi Stroke Berdasarkan Kelainan Patologis dengan Learning Vector Quantization". Jurnal EECCIS Vol.8.
- Daihani, DU., 2001, Komputerisasi Pengambilan Keputusan, PT Elex Media Komputindo Gramedia, Jakarta.
- Kusrini. 2007. Konsep dan Aplikasi Sistem Pendukung Keputusan. Yogyakarta: Andi.
- Kusumadewi S, 2009. "Klasifikasi Status Gizi Menggunakan Naive Bayesian Classification". CommIT, Vol. 3 No. 1 Mei 2009, hlm. 6 - 11.
- Kusumadewi, S, 2003, Artificial Intelligence Teknik dan Aplikasinya, Graha Ilmu, Yogyakarta.
- LaureneFausett, Fundamentals Of Neural Networks. Architectures, Algorithms, And Applications, New Jersey: Prentice-Hall, Inc, 1994.
- Marlinda, Linda. (2004). Sistem Basis Data. Yogyakarta : Andi Offset.
- Nurochman, Sistem Pakar Berbasis Web Untuk Mendiagnosa Penyakit Epilepsi dan Penanganannya Menggunakan Theorema Bayes 2013.
- Sri Kusumadewi, 2009. "Klasifikasi Status Gizi Menggunakan Naive Bayesian Classification". CommIT, Vol. 3 No. 1 Mei 2009, hlm. 6 - 11
- Saptono, 2016 text classification using naive bayes updateable algorithm in sbmptn test questions evaluasi klasifikasi "TELEMATIKA, Vol. 13, No. 02, JULI, 2016, Pp. 123 - 133 ISSN 1829-667X.
- T. Sujoto, S.Si., M.M.Kom, 2013, Kecerdasan Buatan, Penerbit ANDI yogjakarta.
- Wahyudi Setiawan dan Sofie Ratnasari. 2014. "Sistem Pakar Diagnosis Penyakit Mata Menggunakan Naïvebayes Classifier". ISSN : 2407 - 1846,
- Warnia Nengsih dan Ibnu Surya, Analisa Kelayakan Pembukaan Cabang Baru Bisnis Usaha dengan Menggunakan Naive Bayes Classification, Jurnal Teknik Elektro dan Komputer Vol. 2, No. 2, Oktober 2014, 217-224