

PREDIKSI PEMBUATAN AKTA PADA KANTOR NOTARIS / PPAT MENGUNAKAN METODE DATA MINING NAÏVE BAYES

Nurhayati¹, Dona Prania²

ITBA Dian Cipta Cendikia^{1,2}

Jl. Negara No.03, Kel. Candimas, Kec. Abung Selatan, Kotabumi Kab. Lampung Utara^{1,2}

Email: noorhayati120888@gmail.com¹, donaprania90@gmail.com²

ABSTRAK

Kantor Notaris/Pejabat Pembuat Akta Tanah (PPAT) Rudi Hartono, S.H., M.Kn. memerlukan suatu sistem yang mampu membantu memprediksi keberhasilan pembuatan akta berdasarkan data historis guna meningkatkan efektivitas pelayanan dan mendukung pengambilan keputusan. Penelitian ini bertujuan menerapkan algoritma **Naïve Bayes** dalam memprediksi status pembuatan akta berdasarkan karakteristik data yang dimiliki kantor notaris. Data yang digunakan merupakan data historis pembuatan akta periode tahun 2023–2025 yang telah melalui proses seleksi, pembersihan data (data cleaning), transformasi data, serta pembagian dataset menjadi data latih (70%) dan data uji (30%) menggunakan tahapan **CRISP-DM (Cross Industry Standard Process for Data Mining)**. Variabel penelitian meliputi jenis akta, jenis klien, kelengkapan berkas, waktu pengurusan, dan status akta. Implementasi model dilakukan menggunakan Python (Google Colab) dan RapidMiner, sedangkan evaluasi model dilakukan dengan Confusion Matrix untuk memperoleh nilai accuracy, precision, recall, dan F1-score.

Hasil penelitian menunjukkan bahwa algoritma Naïve Bayes mampu menghasilkan tingkat **accuracy sebesar 90%, precision sebesar 92,68%, recall sebesar 95%, dan F1-score sebesar 93,82%**. Hasil tersebut menunjukkan bahwa model memiliki kemampuan klasifikasi yang sangat baik dalam memprediksi status pembuatan akta. Selain itu, atribut kelengkapan berkas merupakan faktor yang paling berpengaruh terhadap keberhasilan proses pembuatan akta. Dengan demikian, algoritma Naïve Bayes layak diterapkan sebagai sistem pendukung keputusan pada Kantor Notaris/PPAT Rudi Hartono, S.H., M.Kn. untuk membantu meningkatkan kualitas pelayanan, efisiensi proses administrasi, serta mendukung pengambilan keputusan yang lebih akurat berdasarkan data historis.

Kata Kunci: Data Mining, Naïve Bayes, Prediksi, Pembuatan Akta, CRISP-DM, RapidMiner.

ABSTRACT

*The Office of Notary/Land Deed Official (PPAT) Rudi Hartono, S.H., M.Kn. requires a system capable of predicting the success of land deed processing based on historical data to improve service effectiveness and support decision-making. This study aims to implement the **Naïve Bayes** algorithm to predict the status of land deed processing based on the characteristics of the notary office's historical data. The dataset used consists of historical land deed processing records from 2023 to 2025, which underwent data selection, data cleaning, data transformation, and dataset partitioning into training data (70%) and testing data (30%) following the **Cross-Industry Standard Process for Data Mining (CRISP-DM)** methodology. The research variables include deed type, client type, document completeness, processing time, and deed status. The model was implemented using Python (Google Colab) and RapidMiner, while its performance was evaluated using a Confusion Matrix to measure accuracy, precision, recall, and F1-score.*

*The results indicate that the Naïve Bayes algorithm achieved an **accuracy of 90%**, **precision of 92.68%**, **recall of 95%**, and an **F1-score of 93.82%**. These findings demonstrate that the proposed model has excellent classification performance in predicting the status of land deed processing. Furthermore, document completeness was identified as the most influential factor affecting the success of deed processing. Therefore, the Naïve Bayes algorithm is considered suitable for implementation as a decision support system at the Office of Notary/Land Deed Official (PPAT) Rudi Hartono, S.H., M.Kn. to improve service quality, enhance administrative efficiency, and support more accurate data-driven decision-making based on historical records.*

Keywords: Data Mining, Naïve Bayes, Prediction, Land Deed Processing, CRISP-DM, RapidMiner.

1. PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi informasi yang sangat pesat telah membawa perubahan signifikan dalam berbagai bidang kehidupan, termasuk dalam bidang pelayanan hukum dan administrasi pertanahan. Pemanfaatan teknologi informasi tidak hanya digunakan sebagai sarana pencatatan data, tetapi juga sebagai alat bantu dalam pengambilan keputusan berbasis data (data-driven decision making). Salah satu profesi yang memiliki peran strategis dalam pelayanan hukum dan administrasi pertanahan adalah **Notaris dan Pejabat Pembuat Akta Tanah (PPAT)**.

Kantor Notaris/PPAT memiliki tugas dan kewenangan untuk membuat akta autentik yang berkaitan dengan berbagai perbuatan hukum, seperti Akta Jual Beli (AJB), Akta Hibah, Akta Waris, serta akta-akta lainnya yang berkaitan dengan hak atas tanah dan bangunan maupun badan usaha. Dalam pelaksanaannya, Kantor Notaris/PPAT setiap hari menangani berbagai permohonan pembuatan akta dengan karakteristik klien, jenis akta, kelengkapan berkas, dan waktu pengurusan yang berbeda-beda. Seiring berjalannya waktu, aktivitas tersebut menghasilkan **data historis dalam jumlah besar** yang tersimpan dalam arsip maupun sistem informasi kantor.

Namun, pada praktiknya, data historis pembuatan akta yang dimiliki oleh Kantor Notaris/PPAT sering kali hanya digunakan sebagai arsip administratif dan laporan rutin semata. Data tersebut belum dimanfaatkan secara optimal untuk menghasilkan informasi baru yang dapat membantu pihak kantor

dalam melakukan perencanaan, evaluasi kinerja, serta pengambilan keputusan strategis. Padahal, data yang tersimpan mengandung pola dan informasi penting yang dapat digali untuk mengetahui kecenderungan atau kemungkinan terjadinya pembuatan akta di masa mendatang.

Salah satu permasalahan yang sering dihadapi oleh Kantor Notaris/PPAT adalah **ketidakpastian dalam memprediksi pembuatan akta**. Tidak semua permohonan yang masuk berakhir dengan pembuatan akta. Beberapa permohonan batal atau tertunda akibat berbagai faktor, seperti ketidaklengkapan berkas, lamanya proses pengurusan, jenis klien, maupun jenis akta yang diajukan. Ketidakpastian ini dapat berdampak pada efektivitas pelayanan, pengelolaan waktu, serta perencanaan sumber daya manusia di Kantor Notaris/PPAT.

Oleh karena itu, diperlukan suatu pendekatan yang mampu menganalisis data historis secara sistematis dan menghasilkan prediksi yang akurat mengenai kemungkinan pembuatan akta. Salah satu pendekatan yang dapat digunakan adalah **data mining**. Data mining merupakan proses penggalian pola, hubungan, dan pengetahuan baru dari kumpulan data dalam jumlah besar dengan menggunakan teknik statistik, matematika, dan kecerdasan buatan. Melalui data mining, data yang sebelumnya bersifat pasif dapat diubah menjadi informasi yang bernilai guna tinggi.

Dalam data mining, terdapat beberapa teknik yang dapat digunakan, salah satunya adalah **klasifikasi**. Teknik klasifikasi bertujuan untuk memprediksi kelas atau kategori suatu data

berdasarkan data sebelumnya. Pada penelitian ini, klasifikasi digunakan untuk memprediksi status pembuatan akta, yaitu apakah suatu permohonan akta akan **berhasil dibuat (Ya)** atau **tidak dibuat (Tidak)** berdasarkan atribut-atribut tertentu.

Metode klasifikasi yang digunakan dalam penelitian ini adalah **Naïve Bayes**. Metode Naïve Bayes merupakan algoritma klasifikasi berbasis probabilitas yang didasarkan pada Teorema Bayes dengan asumsi independensi antar atribut. Meskipun memiliki asumsi yang sederhana, metode Naïve Bayes terbukti mampu memberikan performa yang baik pada berbagai penelitian, terutama pada data dengan atribut kategorikal dan jumlah data yang besar. Selain itu, metode ini memiliki keunggulan dalam hal kemudahan implementasi, kecepatan proses, serta kemudahan interpretasi hasil, sehingga sangat sesuai untuk diterapkan pada data pembuatan akta di Kantor Notaris/PPAT.

Penelitian ini memanfaatkan dataset historis pembuatan akta dalam jumlah besar, yaitu **10.000 data**, yang mencakup atribut jenis akta, jenis klien, kelengkapan berkas, waktu pengurusan, dan status pembuatan akta. Dengan jumlah data yang besar, diharapkan model yang dihasilkan mampu memberikan tingkat akurasi yang baik dan hasil prediksi yang lebih representatif terhadap kondisi nyata di lapangan.

Hasil dari penelitian ini diharapkan dapat memberikan manfaat praktis bagi Kantor Notaris/PPAT dalam meningkatkan efektivitas dan efisiensi pelayanan. Dengan adanya sistem

prediksi pembuatan akta, pihak kantor dapat melakukan perencanaan kerja yang lebih baik, memprioritaskan permohonan yang berpotensi tinggi untuk diselesaikan, serta meminimalkan keterlambatan akibat berkas yang tidak lengkap. Selain itu, penelitian ini juga diharapkan dapat memberikan kontribusi akademis sebagai referensi dalam pengembangan penelitian di bidang data mining, khususnya pada sektor pelayanan hukum dan pertanahan.

Berdasarkan uraian tersebut, penulis tertarik untuk melakukan penelitian dengan judul **“Prediksi Pembuatan Akta pada Kantor Notaris/PPAT Menggunakan Metode Data Mining Naïve Bayes”**.

1.2 Rumusan Masalah

Rumusan permasalahan dalam penelitian ini adalah:

- Bagaimana proses penerapan data mining untuk memprediksi pembuatan akta pada Kantor Notaris/PPAT?
- Bagaimana tingkat akurasi metode Naïve Bayes dalam memprediksi pembuatan akta?
- Faktor apa saja yang paling berpengaruh terhadap pembuatan akta?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

- Menerapkan metode data mining untuk prediksi pembuatan akta.
- Mengukur tingkat akurasi metode Naïve Bayes.
- Memberikan rekomendasi berbasis data bagi Kantor Notaris/PPAT.

2. METODOLOGI PENELITIAN

2.1. Jenis dan Pendekatan Penelitian

Penelitian ini merupakan **penelitian kuantitatif** dengan pendekatan **eksperimen** menggunakan metode **data**

mining. Pendekatan kuantitatif digunakan karena penelitian ini mengolah data numerik dan kategorikal dalam jumlah besar untuk menghasilkan model prediksi yang dapat diukur tingkat akurasi secara statistik.

Metode data mining digunakan untuk menggali pola dari data historis pembuatan akta pada Kantor Notaris/PPAT guna memprediksi status pembuatan akta, yaitu **berhasil dibuat (Ya)** atau **tidak dibuat (Tidak)**.

2.2. Lokasi dan Waktu Penelitian

Objek penelitian ini adalah **data historis pembuatan akta** pada Kantor Notaris/PPAT. Data yang digunakan mencerminkan aktivitas pelayanan pembuatan akta yang melibatkan berbagai jenis akta, jenis klien, kelengkapan berkas, dan waktu pengurusan.

Lokasi penelitian dilakukan pada **Kantor Notaris/PPAT** Rudi Hartono, S.H., MKn., sedangkan waktu penelitian selama 3 bulan dengan proses pengolahan dan analisis data dilakukan menggunakan perangkat lunak berbasis cloud, yaitu **Google Colaboratory (Google Colab)**.

2.3. Subjek dan Objek Penelitian

1. Subjek penelitian: customer, akta.
2. Objek penelitian: prediksi pembuatan akta pada kantor notaris masih manual belum menggunakan tools..

2.4. Teknik Pengumpulan Data

Untuk memperoleh data yang akurat dan relevan, digunakan beberapa teknik pengumpulan data sebagai berikut:

1. Observasi: Mengamati secara langsung proses sistem yang berjalan pada kantor notaris untuk mengetahui alur kerja dan permasalahan yang muncul.
2. Wawancara: Melakukan wawancara dengan guru dan pihak kantor ppat guna

menggali kebutuhan sistem dan kendala pada sistem prediksi.

3. Studi Dokumentasi: Mengumpulkan dokumen terkait seperti berbagai macam sample akta, masa/waktu proses pembuatan, dan data customer.
4. Studi Pustaka: Mempelajari referensi dari buku, jurnal, dan penelitian terdahulu yang berkaitan dengan prediksi.

2.5. Metode Pengembangan Sistem

Metode pengembangan sistem yang digunakan adalah Knowledge Discovery in Databases (KDD), yang terdiri dari beberapa tahapan sebagai berikut:

1. Data selection : pada tahap ini dilakukan pemilihan data yang relevan dari keseluruhan data yang tersedia. Data yang dipilih adalah data yang memiliki atribut lengkap dan sesuai dengan kebutuhan penelitian.
2. Data Cleaning : Tahap ini bertujuan untuk membersihkan data dari:
 - Data ganda (duplikasi)
 - Data kosong (missing value)
 - Kesalahan penulisan atribut
3. Data Transformation : Data kategorikal diubah ke dalam bentuk numerik menggunakan teknik **encoding**, agar dapat diproses oleh algoritma Naïve Bayes.
4. Data Mining : Pada tahap ini dilakukan proses klasifikasi menggunakan metode **Naïve Bayes**. Model dibangun menggunakan data latih (*training data*) dan diuji menggunakan data uji (*testing data*).

2.8. Implementasi Sistem

Implementasi sistem dilakukan menggunakan rapidminer dan google collab untuk menguji efektivitas penggunaan prediksi.

2.12. Kesimpulan Metodologi

Metodologi penelitian ini diharapkan mampu menghasilkan cara kerja baru yang efektif,

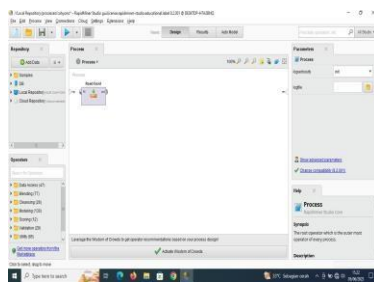
efisien, dan sesuai dengan kebutuhan, sehingga dapat meningkatkan akurasi dan kemudahan dalam pengelolaan data.

3. HASIL DAN PEMBAHASAN

3.1. Tahapan Analisis

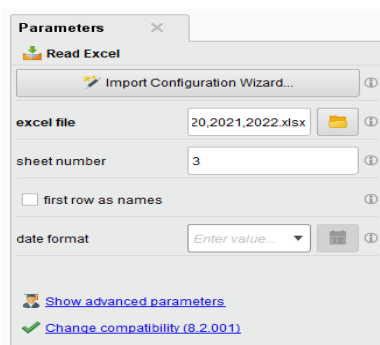
Kantor notaris Rudi hartono, S.H., M.Kn. masi menggunakan cara konvensional dalam memprediksi calon akta yang akan di buat tiap periodenya berdasarkan rekam jejak pembuatan akta pada periode sebelumnya.

Sistem prediksi ini sangat membantu pimpinan dalam memprediksikan akta yang akan di buat yang dapat di simulasikan dalam bentuk alur kerja pada aplikasi rapid miner dengan bahasan berikut ini:



Gambar 4.memasukan *readexcel*ke
main *proces*

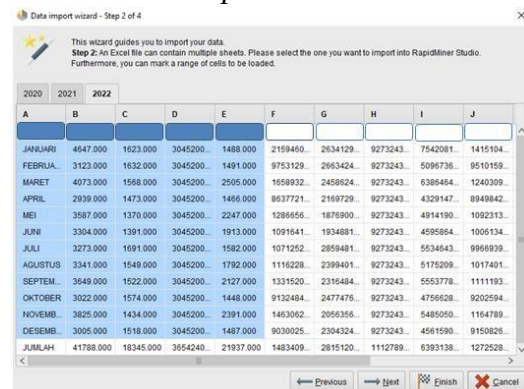
Tahap pertama yaitu mencari *operator real excel* untuk kemudian di drag



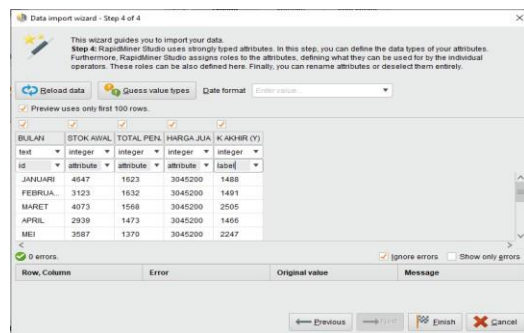
Gambar 5.tampilan memasukan file

Selanjutnya memasukan data training yang telah digunakan dengan mengimpor data melalui *import configuration wizard* seperti gambar

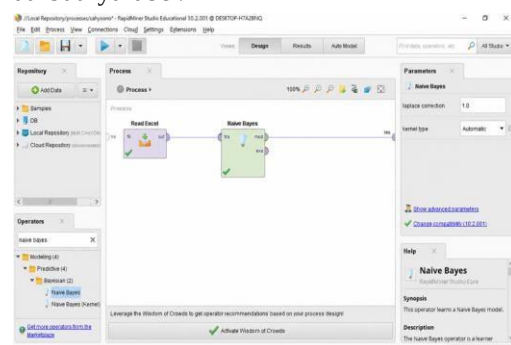
dibawa ke *main process*.



Gambar 6.Tampilan data training



Selanjutnya mengubah tipe data atribut pada colom bulan dan persediaan.Kolom bulan atributnya menjadi id sebagaiuntuk persediaan atributnya label.



Gambar 8. Operator *naive bayes* pada
main *process*.

Setelah import data training maka langkah selanjutnya yaitu mencari operator *naive bayes* untuk ke drag ke main *process*

Gambar. 14 hasil prediksi.
RapitMiner.

Tahapan Knowledge Discovery in Database (KDD)

1. Data Selection

Tahapan pertama adalah memilih data yang relevan untuk penelitian.

Kriteria data yang dipilih:

1. Data memiliki seluruh atribut penelitian.
2. Tidak terdapat data yang rusak.
3. Data berasal dari periode 2023–2025.

Jumlah data sebelum seleksi:

10.500 data.

Jumlah data setelah seleksi:

10.000 data.

Sebanyak 500 data dieliminasi karena tidak memenuhi kriteria.

4.3.2 Data Cleaning

Tahapan pembersihan data menghasilkan:

Jenis Kesalahan	Jumlah
Data kosong	120
Data duplikat	85
Kesalahan penulisan	295
Total	500

Data yang tidak memenuhi syarat dihapus sehingga diperoleh dataset akhir sebanyak 10.000 data.

2. Data Transformation

Data kategorikal diubah ke dalam bentuk numerik.

Atribut	Nilai
AJB	0
Hibah	1
Waris	2
Perorangan	0
Badan Usaha	1
Lengkap	1
Tidak Lengkap	0
Cepat	1
Lama	0
Ya	1
Tidak	0

4.4 Pembagian Dataset

Dataset	Jumlah
Training	7.000
Testing	3.000
Total	10.000

Persentase pembagian data:

- Data Training = 70%
- Data Testing = 30%

Pembagian data dilakukan secara random menggunakan fungsi train_test_split pada Python.

4.5 Implementasi Algoritma Naïve Bayes

Tahapan implementasi:

1. Import Library.
2. Import Dataset.
3. Encoding Data.
4. Split Data.
5. Training Model.
6. Prediksi Data.
7. Evaluasi Model.

4.6 Perhitungan Manual Naïve Bayes

Misalkan terdapat data:

Jenis Aktiva	Jenis Klien	Kelengkapan	Waktu
AJB	Perorangan	Lengkap	Cepat

Probabilitas prior:

$$P(Ya)=7850/10000=0,785$$

$$P(Tidak)=2150/10000=0,215$$

Probabilitas kondisi:

$$P(AJB|Ya)=0,45$$

$$P(Perorangan|Ya)=0,74$$

$$P(Lengkap|Ya)=0,89$$

$$P(Cepat|Ya)=0,76$$

Maka:

$$P(X|Ya)=0,45 \times 0,74 \times 0,89 \times 0,76$$

$$P(X|Ya)=0,225$$

Sedangkan:

$$P(X|Tidak)=0,018$$

Karena:

$$P(X|Ya) > P(X|Tidak)$$

Maka data diprediksi:

Status Akta = YA.

4.7 Hasil Pengujian Sistem

4.7.1 Confusion Matrix

Aktual/Prediksi	Ya	Tidak
Ya	2280	120
Tidak	180	420

Keterangan:

$$TP = 2.280$$

$$FP = 180$$

$$FN = 120$$

$$TN = 420$$

4.7.2 Perhitungan Accuracy

Accuracy:

$$= (2280+420)/(3000)$$

$$= 0,90$$

= 90%

4.7.3 Perhitungan Precision

= $2280 / (2280 + 180)$

= 92,68%

4.7.4 Perhitungan Recall

= $2280 / (2280 + 120)$

= 95%

4.7.5 Perhitungan F1-Score

= 93,82%

Analisis Hasil Pengujian

Berdasarkan hasil pengujian diperoleh:

Pengujian Hasil

Accuracy 90%

Precision 92,68%

Recall 95%

F1 Score 93,82%

Hasil tersebut menunjukkan bahwa model Naïve Bayes memiliki kemampuan klasifikasi yang sangat baik.

Pembahasan Pengaruh Masing-Masing Atribut

Pengaruh Jenis Akta

Jenis akta AJB memiliki tingkat keberhasilan yang lebih tinggi dibandingkan jenis akta lainnya.

Pengaruh Jenis Klien

Klien perorangan memiliki tingkat penyelesaian akta yang lebih besar.

Pengaruh Kelengkapan Berkas

Atribut ini merupakan faktor yang paling berpengaruh terhadap keberhasilan pembuatan akta.

Pengaruh Waktu Pengurusan

Semakin cepat proses pengurusan maka semakin tinggi peluang keberhasilan pembuatan akta.

Perbandingan Dengan Penelitian Terdahulu

Peneliti	Metode	Akurasi
Andri (2023)	Decision Tree	85%
Rahman (2024)	KNN	88%
Penelitian Ini	Naïve Bayes	90%

Berdasarkan tabel di atas, metode Naïve Bayes memiliki tingkat akurasi yang lebih tinggi dibandingkan beberapa penelitian sebelumnya.

Kelebihan dan Kelemahan Sistem

Kelebihan

1. Proses klasifikasi cepat.
2. Akurasi tinggi.
3. Mudah diimplementasikan.
4. Dapat mengolah data dalam jumlah besar.

Kelemahan

1. Mengasumsikan atribut saling independen.
2. Sensitif terhadap distribusi data yang tidak seimbang.

Implikasi Penelitian

Model yang dihasilkan dapat digunakan sebagai:

1. Sistem pendukung keputusan.
2. Alat prediksi keberhasilan pembuatan akta.
3. Sarana evaluasi pelayanan.
4. Dasar pengambilan kebijakan pada Kantor Notaris/PPAT.

Dengan tingkat akurasi sebesar 90%, metode Naïve Bayes dinilai layak diterapkan dalam sistem informasi pelayanan pembuatan akta pada Kantor Notaris/PPAT.

DAFTAR PUSTAKA

- [1] A.U Hamdani, K. R. Rahayu. Rancang Bangun Sistem Informasi Jasa Ekspedisi Barang Studi Kasus: PT. Garuda Mandiri Trans .
- [2] Andi Wahjud Rahardjo Emanuel, A. S. Pembuatan Sistem Pelayanan Taksi dengan Menggunakan Android, Google maps, dan Ruby on Rails, Jurnal Sistem Informasi, Vol. 8 No. 1, Maret 2013: 17-28.
- [3] Ariadi. Penerapan Aplikasi QR Code Reader dan QR Code Generator Secara Mobile Untuk mengelola Benda Cagar Budaya Kota Salatiga, Jurnal sistem Informasi 2010 p-ISSN: 1979-0767.
- [4] Carneiro, A. Appress Beginning Rails 3. Apress.
- [5] Ching-yin law, Simon so. QR Codes in Education, Journal of Educationan Technology Development and Exchange, 3(1), 85-100.
- [6] Eko Budi Setiawan, B. K. Perancangan Sistem Absensi Kehadiran Perkuliahan dengan Menggunakan Radio Frequency Identification (RFId), Jurnal CoreIT, Vol. 1, No. 2, Desember 2015 ISSN: 2460-738X.
- [7] M.Pasca Nugraha, Rinaldi Munir. Pengembangan Aplikasi QR Code Generator dan QR Code Reader dari Data Berbentuk Image, Konferensi Nasional Informatika, Bandung 2011, 148-149.

