

## PENERAPAN DATA MINING UNTUK KLASIFIKASI PRODUK (OK DAN SCRAP) PADA INDUSTRI BAN MENGGUNAKAN METODE NAIVE BAYES

Donny Maulana, Yusup Solikhun

Program Studi Teknik Informatika Sekolah Tinggi Teknologi Pelita Bangsa  
donny.maulana@pelitabangsa.ac.id

Disetujui, 05 Maret 2018

### Abstrak

Masih banyaknya ketidaksesuaian terhadap kualitas ban secara visual inspeksi. Visual inspeksi yang dilakukan sebelum proses pengiriman masih banyak ditemukan miss inspection tidak sesuai standar kualitas yang sudah ditentukan. Untuk itu diperlukan suatu sistem yang bisa mengklasifikasikan produk baik produk OK atau Scrap. Dari kriteria yang ditetapkan membutuhkan algoritma yang tepat untuk melakukan perhitungan agar hasil yang diberikan lebih akurat. Algoritma *Naïve Bayes* merupakan metode untuk klasifikasi dengan menggunakan teori probabilitas yang memiliki tingkat akurasi tinggi. Pengujian algoritma *Naïve Bayes* menggunakan *tools Rapid Miner* yang menghasilkan tingkat akurasi sebesar 96% dari 200 data yang diberikan. Algoritma ini tepat digunakan untuk klasifikasi produk. Terdapat 2 class yang dibutuhkan yaitu *Ok* dan *Scrap*.

Kata kunci: Klasifikasi, *Naïve Bayes*, *Rapid Miner*, produk.

### Abstract

*There are still many incompatibilities with tire quality visually inspection. Visual inspections carried out before the shipping process are still found that miss inspection does not meet the specified quality standards. For that we need a system that can classify products either OK or Scrap products. From the established criteria requires the right algorithm to do calculations so that the results given are more accurate. Naïve Bayes algorithm is a method for classification using probability theory that has a high degree of accuracy. Testing the Naïve Bayes algorithm uses Rapid Miner tools that produce an accuracy rate of 96% of the 200 data provided. This algorithm is appropriate for product classification. There are 2 classes needed, namely Ok and Scrap.*

*Keywords: Classification, Naïve Bayes, Rapid Miner, products.*

### 1. Pendahuluan

Semakin pesatnya perkembangan ilmu pengetahuan dan teknologi yang semakin canggih, sejalan dengan kemajuan zaman maka penggunaan komputer semakin meningkat dan perannya sangat penting dalam membantu

pekerjaan manusia. Persaingan usaha yang semakin ketat ini membuat manajemen perusahaan harus mengatur kebijakan terhadap aset yang dimiliki oleh perusahaan terutama aset tetap yang merupakan elemen penting dalam menunjang aktifitas operasi perusahaan. Salah satu aset tetap yang sangat vital dalam kegiatan operasi suatu perusahaan adalah kualitas.

Pengendalian kualitas produk dalam suatu proses produksi adalah faktor yang sangat penting bagi dunia industri, selain untuk menjaga konsistensi kualitas produk yang dihasilkan sesuai dengan kebutuhan pasar (Sunardi, A. T. P., & Suprianto, E., 2018). Hal ini dilakukan karena pengendalian kualitas yang baik dan dilakukan secara kontinu akan dapat mendeteksi ketidaksesuaian secara cepat, sehingga dapat segera dilakukan tindakan antisipasinya (Saripudain, A. A., & Satar, M., 2018). Pengendalian kualitas ini jika tidak dilakukan dengan tepat akan menyebabkan banyak bahan, tenaga dan waktu yang terbuang, sehingga muncul pemikiran untuk menciptakan sistem yang dapat mencegah timbulnya masalah mengenai kualitas (Wisnubroto, P., & Rukmana, A., 2015).

Klasifikasi produk sangat penting untuk menentukan kualitas produk yang baik. Untuk mengklasifikasikan data dibutuhkan

algoritma untuk menyelesaikannya, salah satunya adalah algoritma *Naïve Bayes*. *Naïve Bayes Classification* merupakan salah satu metode *machine learning* yang menggunakan perhitungan probabilitas. Konsep dasar yang digunakan oleh *Naïve Bayes* adalah teorema *bayes* yaitu teorema dalam statistika untuk menghitung peluang, *bayes optimal classifier* menghitung peluang dari suatu kelas dari masing-masing kelompok atribut yang ada, dan menentukan kelas mana yang paling optimal.

*Naïve Bayes* dapat digunakan untuk meneliti berbagai macam hal, diantaranya adalah prediksi penggunaan listrik rumah tangga, klasifikasi kelulusan mahasiswa, penentuan kelayakan calon Tenaga Kerja Indonesia (TKI), dan lain-lain. Sebelumnya juga telah ada yang melakukan penelitian yang serupa dengan menggunakan metode ini, namun atribut yang digunakan berbeda. Adapun pembeda lainnya adalah penelitian tersebut dilakukan untuk klasifikasi kematangan buah apel manalagi dengan algoritma *naïve bayes* dan ekstraksi fitur citra digital. Penelitian tersebut dilakukan oleh Antonio Ciputra, De Rosal Ignatius Moses Setiadi, Eko Hari Rachmawanto, Ajib Susanto. Vol. 9 No.1 April 2018 ISSN: 2252-4983. Tingkat akurasi yang dihasilkan dari metode *Naïve Bayes* ini sudah cukup tinggi, oleh karena itu dapat disimpulkan metode ini cocok untuk diimplementasikan pada penelitian ini. *Scrap*.

## 2. Tinjauan Studi

Ban adalah peranti yang menutupi *velg* suatu roda. Ban adalah bagian penting dari kendaraan darat, dan digunakan untuk mengurangi getaran yang disebabkan ketidakrataan permukaan jalan, melindungi roda dari aus dan kerusakan, serta memberikan kestabilan antara kendaraan dan tanah untuk meningkatkan percepatan dan mempermudah pergerakan. Sebagian besar ban yang ada sekarang, terutama yang digunakan untuk kendaraan bermotor, diproduksi dari karet sintetik, walaupun dapat juga digunakan dari bahan lain seperti baja.

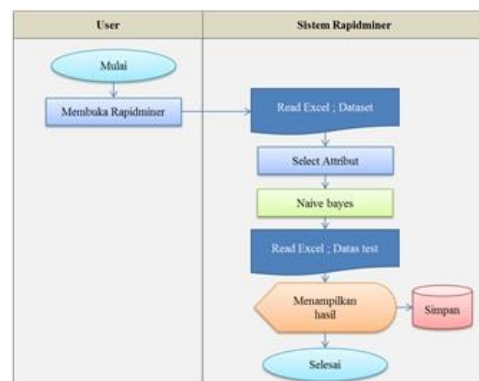
*Naïve Bayes Classifier* disebut juga *Bayesian Classification* adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. *Naïve Bayes Classifier* didasarkan pada teorema Bayes yang memiliki kemampuan klasifikasi serupa dengan *decision tree* dan *neural network*. Selain itu, *Naïve Bayes Classifier* terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar. Bayes merupakan teknik klasifikasi berbasis probabilistic sederhana yang

berdasar pada penerapan teorema Bayes (aturan Bayes) dengan asumsi independensi yang kuat (*Naïve*) dengan kata lain *Naïve Bayes Classifier*. Model yang digunakan adalah model fitur independen. Dalam Bayes terutama *Naïve Bayes Classifier*, maksud independen yang kuat dalam fitur adalah bahwa sebuah fitur pada data tidak berkaitan dengan ada atau tidaknya fitur lain pada data yang sama. (Via, Nugroho, & Syafrizal, 2015)

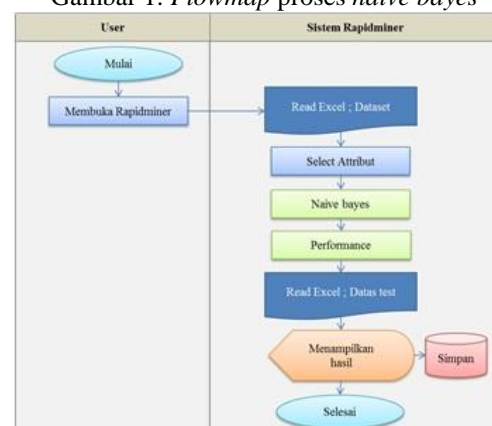
## 3. Kerangka Konsep

Pengujian metode dilakukan untuk mengetahui hasil perhitungan yang dianalisa dan untuk mengetahui apakah fungsi bekerja dengan baik atau tidak. Setelah data dihitung secara manual, kemudian data diuji menggunakan tools *Rapid Miner* untuk memastikan apakah hasil perhitungan manual dengan hasil yang diperoleh *Rapid Miner* sama atau tidak. Sehingga diketahui hasil dari akurasi perhitungan menggunakan metode *naive bayes* *Flowmap* Proses

*Flowmap* proses menunjukkan aliran proses dari mulai sampai dengan hasil yang sudah diklasifikasi menggunakan metode *naive bayes*.



Gambar 1. *Flowmap* proses *naive bayes*



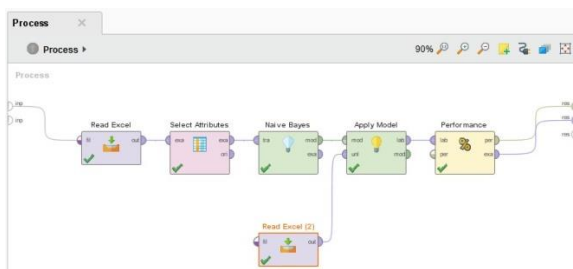
Gambar 2. *Flowmap* proses akurasi *naive bayes*

## 4. Desain Penelitian/Metodologi

Melakukan pengecekan terhadap setiap nilai atribut dan model yang sudah dibangun, kemudian melakukan evaluasi terhadap hasil dengan melakukan analisis *data mining*. Pada tahap ini juga merupakan tahapan dimana dilakukan perbaikan kembali bila terjadi *kejointan*, bisa saja kembali lagi ke tahap yang pertama dan kemudian ke tahap berikutnya dengan tujuan perbaikan sampai sesuai dengan kebutuhan. Variabel penelitian yang akan digunakan sebagai atribut data untuk proses *data mining* ialah tipe ban, jenis ban, nilai, lokasi, tipe *carcass*, kode *pattern*.. Kelas hasil adalah variabel target penelitian yang berisi 2 nilai kelas yaitu “Ok” dan “Scrap”.

**5. Hasil Penelitian**

Proses akurasi dengan *Rapid Miner* menggunakan metode *Naive Bayes* yang digunakan untuk mengklasifikasi data ban pada penelitian ini.



Gambar 3. Proses Data *Training* dan Data *Testing*

a. *Accuracy* / akurasi

Dengan mengetahui jumlah data yang di klasifikasikan secara benar maka dapat diketahui akurasi hasil prediksi yaitu 96% dari hasil data *testing*.

accuracy: 96.00%

	true SCRAP	true OK	class precision
pred. SCRAP	26	2	92.86%
pred OK	0	22	100.00%
class recall	100.00%	91.67%	

Gambar 4. *Accuracy* / Akurasi

Untuk menghitung akurasinya sebagai berikut:

- Jumlah data yang diuji = 50
- Jumlah data yang diprediksi benar “OK” = 22
- Jumlah data yang diprediksi benar “SCRAP” = 26
- Jumlah data yang diprediksi salah “OK” = 2

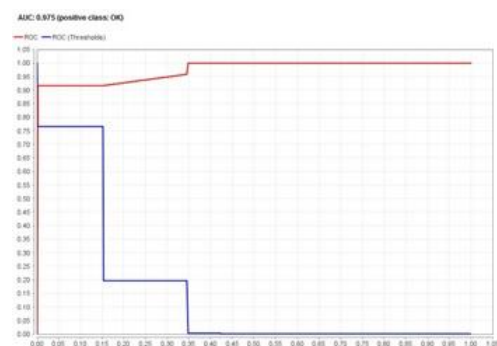
Jumlah data yang diprediksi salah “SCRAP” = 0

$$\begin{aligned} \text{Akurasi} &= \text{Jumlah data yang diprediksi benar} / \text{jumlah data yang diuji} * 100\% \\ &= (48) / (50) * 100\% \\ &= 96\% \end{aligned}$$

$$\begin{aligned} \text{Error} &= \text{Jumlah data yang diprediksi salah} / \text{jumlah data yang diuji} * 100\% \\ &= (2) / (50) * 100\% \\ &= 4\% \end{aligned}$$

Dari perhitungan tersebut dapat disimpulkan bahwa perhitungan dengan menggunakan metode *Naive Bayes* yang dibuat menghasilkan tingkat akurasi tinggi sebesar 96% dan tingkat *error* 4%.

b. Kurva ROC



Gambar 5. Kurva ROC

Kurva ROC digunakan untuk mengekspresikan data. Garis horisontal mewakili nilai *false positive* dan garis vertikal mewakili nilai *true positive*. Dari gambar 4.5 dapat diketahui bahwa nilai *Area Under Curve (AUC)* model algoritma *Naive Bayes* adalah 0.975, hal ini menunjukkan bahwa model algoritma *Naive Bayes* mencapai klasifikasi hampir sempurna.

**6. Kesimpulan**

Berdasarkan hasil penelitian yang telah dilakukan mengenai klasifikasi untuk menentukan ban OK dan *Scrap* dengan metode *Naive Bayes* di PT. Hankook Tire Indonesia, maka dapat diambil beberapa kesimpulan sebagai berikut:

- a. Penerapan *data mining* menggunakan algoritma *Naive Bayes* sangat cocok diterapkan dalam klasifikasi ban OK atau *Scrap*, karena mampu menentukan hasil

dengan baik. Sehingga memudahkan operator dalam menentukan ban tersebut OK atau *Scrap*.

- a. Dari pengujian yang dilakukan dengan membandingkan hasil analisa sistem dengan *data training* pada *tools Rapid miner* menghasilkan tingkat akurasi sebesar 96% dan *error* sebesar 4%, sedangkan evaluasi dengan kurva ROC dengan nilai *Area Under Curve (AUC)* model algoritma *Naïve Bayes* adalah 0.975.
- b. Setelah dilakukan pengujian untuk menentukan hasil klasifikasi ban OK atau *Scrap* dengan menggunakan perhitungan manual dan menggunakan aplikasi *Rapid Miner*, didapatkan hasil prediksi yang sama.

#### Daftar Pustaka

- Ciputra, Antonio dkk, Fakultas Ilmu Komputer (2018). Klasifikasi tingkat kematangan buah apel manalagi dengan algoritma *naive bayes* dan ekstraksi fitur citra digital. *Jurnal SIMETRIS*, 461-472
- Haryati, S., Sudarsono, A., & Suryana, E. (2015). Implementasi Data Mining untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5. *Jurnal Media Infotama*, 11(2), 130–138.
- HM, Jogyanto (2001). Perancangan Sistem Informasi. Yogyakarta: Andi
- Jananto, A. (2013). Algoritma Naïve Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa  $P(H|X)P(X|H)P(H)$ , 18(1), 9–16.
- Montororing, Yuri D.R., Universitas Bhayangkara Jakarta Raya (2018). Model Alokasi Stasiun Inspeksi Dengan Mempertimbangkan Work In Process Dan Produk Rework Dan Scrap. *Jurnal Kajian Ilmiah*
- Prasetyo, (2012). Data Mining Konsep dan Aplikasi Menggunakan Matlab. Penerbit Andi Yogyakarta.
- Ridwan, M., Suyono, H., & Sarosa, M. (2013). Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naïve Bayes Classifier. *Eecis*, 7(1), 59–64. <https://doi.org/10.1038/hdy.2009.180>
- Saleh, A. (2015). Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga, 2(3), 207–217.
- Saripudain, A. A., & Satar, M. (2018). ANALISA PENGENDALIAN KUALITAS PRODUK BRCKET ELECTRK AIR BUS 380 DENGAN METODE SIX SIGMA PADA AREA PROFILE PRESS FORMING DI PT. X. *Jurnal Industri Elektro dan Penerbangan*, 4(3).
- Sunardi, A. T. P., & Suprianto, E. (2018). PENGENDALIAN KUALITAS PRODUK PADA PROSES PRODUKSI RIB A320 DI SHEET METAL FORMING SHOP *Jurnal Industr Elektro dan Penerbangan*, 5(2).
- Tampubolon, K., Saragih, H., & Reza, B. (2013). IMPLEMENTASI DATA MINING ALGORITMA APRIORI PADA SISTEM PERSEDIAAN ALAT- ALAT KESEHATAN, 93–106.
- Via, Y. V., Nugroho, B., & Syafrizal, A. (2015). SISTEM PENDUKUNG KEPUTUSAN KLASIFIKASI TINGKAT KEGANASAN KANKER PAYUDARA DENGAN METODE NAIVE, X, 2–7.
- Wisnubroto, P., & Rukmana, A. (2015). Pengendalian Kualitas Produk dengan Pendekatan Six Sigma dan Analisis Kaizen Serta New Seven Tools sebagai Usaha Pengurangan Ke>0.41an Produk. *Jurnal Teknik Industri*, 8(1), 65-74.