

## ANALISA DATA MINING DENGAN METODE KLASIFIKASI UNTUK PRODUK CACAT PADA PT. SHUANGYING INTERNATIONAL INDONESIA

Arif Siswandi<sup>1</sup>, Aswan S. Sunge<sup>2</sup>, Rani Yuliani Wulandari

Program Studi Teknik Informatika Sekolah Tinggi Teknologi Pelita Bangsa

<sup>1</sup>arif.siswandi@pelitabangsa.ac.id, <sup>2</sup>aswan.sunge@pelitabangsa.ac.id

Disetujui, 05 Maret 2018

### Abstrak

PT. Shuangying International Indonesia adalah perusahaan dari negara China pembuat produk plastik otomotif (komponen interior dan eksterior mobil) yang berdiri pada tahun 2000 di negara China, sedangkan di Indonesia mulai beroperasi pada 23 maret 2016. Perkembangan produk otomotif di Indonesia sangat berkembang, hal itu terlihat banyaknya jenis kendaraan otomotif dari luar negeri yang memproduksi dan berbisnis di Indonesia. Pada perusahaan terdapat sebuah bagian pemeriksaan pada Departemen Quality yang bertugas untuk memeriksa produk sebelum dikirim ke pelanggan. Banyaknya produk cacat pada saat proses produksi terbuang karena tidak memiliki pengambilan keputusan secara klasifikasi untuk produk cacat yang bisa membantu mempermudah proses pemeriksaan. Tujuan penelitian ini adalah untuk menganalisa data mining dengan metode klasifikasi untuk produk cacat pada PT. Shuangying International Indonesia dengan algoritma C4.5. Dalam penelitian ini, hasil yang didapatkan dari beberapa atribut menghasilkan sebab-akibat produk cacat dalam mengklasifikasikan menjadi produk OK dan NOK. Penelitian ini diharapkan dapat membantu pihak perusahaan untuk meningkatkan laba perusahaan tersebut. Metode klasifikasi yang digunakan dalam data mining adalah Decision Tree karena dapat menghasilkan output sesuai dengan aturan yang ada, dapat menyajikan data produk cacat. Penelitian menggunakan Algoritma C4.5 untuk menghasilkan aturan klasifikasi produk cacat dari data 90 data training diperoleh produk OK (37), NOK (57) dan hasil akurasi sebesar 95.56%, yang menunjukkan bahwa algoritma C4.5 cocok digunakan untuk mengukur produk cacat pada suatu perusahaan.

Kata kunci: algoritma C4.5, data mining, decision tree, produk cacat.

### Abstract

*PT. Shuangying International Indonesia is a Chinese company that manufactures automotive plastic products (car interior and exterior components) which was established in 2000 in China, while in Indonesia it started operating on March 23, 2016. The development of automotive products in Indonesia is very growing, it can be seen that there are many types of automotive vehicles from abroad that produce and do business in Indonesia. In the company there is an inspection section in the Quality Department which is in charge of checking products before they are sent to customers. The number of defective products during the production process is wasted because it does not have a classification decision for defective products that can help simplify the inspection process. The purpose of this study was to analyze data mining with classification methods for defective products at PT. Shuangying*

*International Indonesia with the C4.5 algorithm. In this study, the results obtained from several attributes resulted in a cause and effect of defective products in classifying OK and NOK products. This research is expected to help the company to increase the company's profit. The classification method used in data mining is the Decision Tree because it can produce output in accordance with existing rules and can present defective product data. The study used the C4.5 Algorithm to produce defective product classification rules from 90 training data obtained by OK (37), NOK (57) products and an accuracy of 95.56%, which shows that the C4.5 algorithm is suitable for measuring defective products in a company.*

*Keywords: C4.5 algorithm, data mining, decision tree, defective product.*

## 1. Pendahuluan

Perkembangan produk otomotif di Indonesia sangat berkembang, hal itu terlihat banyaknya jenis kendaraan otomotif dari luar negeri yang berproduksi dan berbisnis di Indonesia. Seiring perkembangan tersebut maka dibutuhkan pula perkembangan teknologi informasi untuk menembus batas-batas ilmu pengetahuan dan menyerbu pelosok ke seluruh dunia. Revolusi informasi dan komunikasi zaman ini melahirkan peradaban baru yaitu tidak dibatasi oleh waktu, informasi dapat kita peroleh kapan saja dan dimana saja. Begitu pula halnya dalam dunia bisnis, kemajuan teknologi informasi tersebut mengakibatkan perubahan yang signifikan dengan telah banyaknya mengubah dan mempermudah sistem kerja para pelaku bisnis.

Setiap aktivitas dan kegiatan dalam kehidupan berbisnis, komputer dijadikan sebagai suatu perangkat pengolahan data elektronik untuk menghasilkan data informasi dengan cepat dan mudah. Informasi merupakan salah satu jalan untuk mengambil keputusan terhadap apa yang akan dilakukan pada tahap selanjutnya. Sistem informasi yang baik akan meningkatkan efisiensi dan efektivitas kerja.

Dengan banyaknya produk otomotif maka mutu dijadikan sebagai landasan utama. Apapun jenis produk suatu perusahaan tentunya membutuhkan jasa Quality Assurance atau Quality Control dalam suatu divisi. Walaupun keduanya memiliki definisi yang berbeda tetapi sama-sama memiliki acuan pada mutu. Cukup disebut

Departemen Quality didalamnya ada kegiatan pemeriksaan dan merancang jaminan produk bahwa sudah terpenuhi persyaratan mutu tersebut. Secara strategi pun kualitas atau mutu didefinisikan sebagai segala sesuatu yang mampu memenuhi keinginan atau kebutuhan pelanggan (meeting the needs of customer). Keunggulan suatu produk terukur melalui tingkat kepuasan pelanggan maka, diperlukan juga perkembangan dunia teknologi informasi untuk kebutuhan akan data dan informasi berbasis komputer. Data dan informasi sangat dibutuhkan suatu perusahaan yang berskala besar, sedang, dan kecil. Sistem informasi terkomputerisasi terbukti mampu mempercepat kinerja, sehingga meningkatkan keuntungan.

Teknologi data sistem informasi merupakan sebuah teknologi interaktif yang dapat digunakan sebagai sarana informasi yang dapat di akses oleh semua pihak, baik pihak umum maupun pihak yang mempunyai akses. Sebuah perusahaan menggunakan atau menerapkan Teknologi Informasi untuk memudahkan berlangsungnya proses kerja baik untuk semua karyawan, maupun karyawan administrasi dalam hal pengaksesan data dan informasi.

PT. Shuangying International Indonesia adalah sebuah perusahaan manufaktur asal China yang bergerak di bidang pembuatan komponen plastik (interior dan exterior) otomotif untuk kendaraan mobil yang beroperasi di kawasan industri Grendland Deltamas Cikarang Pusat. Pada perusahaan ini terdapat sebuah bagian pemeriksaan pada Departemen Quality yang bertugas untuk memeriksa produk sebelum dikirim ke pelanggan. Akan tetapi pada proses pemeriksaan, inspector mengalami kesulitan dalam prosesnya. Banyaknya produk cacat pada saat proses produksi terbuang karena tidak memiliki pengambilan keputusan secara klasifikasi produk cacat yang bisa membantu mempermudah proses pemeriksaan, sehingga inspector kesulitan dalam mencari limit produk cacat ataupun membuat laporan pemeriksaan setiap harinya. Selain itu lamanya dalam proses pemeriksaan bisa mengurangi hasil produksi dan tentunya hal ini bisa menjadi kerugian untuk perusahaan. Penyediaan data yang cepat dan akurat juga menjadi kendala utama disebabkan ketiadaan sistem informasi yang akurat. Hal tersebut menyulitkan perusahaan dalam mengontrol hasil produksi setiap harinya. Dari beberapa alasan tersebut maka diperlukan sebuah akurasi yang diharapkan mampu mengatasi masalah-masalah yang selama ini menjadi kendala dan mempermudah karyawan untuk menyusun data.

Melihat dari permasalahan, khususnya bagian pemeriksaan pada Departemen Quality penulis bermaksud untuk memberi solusi agar masalah-masalah tersebut dapat terminimalisir. Maka dari itu diperlukan sebuah akurasi yang diharapkan mampu mengatasi masalah-masalah yang selama ini menjadi kendala dan mempermudah pemeriksaan produk cacat sehingga dapat memberikan solusi dari permasalahan yang dihadapi oleh pihak perusahaan.

## 2. Tinjauan Studi

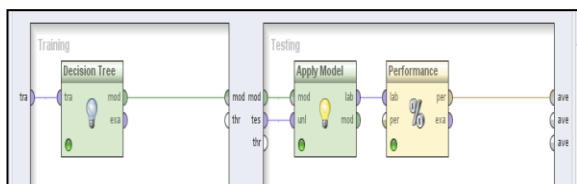
Dari penelitian yang berkaitan dengan penelitian ini terutama dalam hal klasifikasi produk cacat, yaitu:

- a. Prediksi Kompetensi Karyawan Menggunakan Algoritma C4.5 (Studi Kasus : PT. Hankook Tire Indonesia) (Sunge, 2018). Penelitian ini membahas kompetensi karyawan menggunakan model eksperimen data diperoleh dari data sekunder yang diperoleh dari data PT Hankook Tire Indonesia di Bekasi yang datanya dikumpulkan berjumlah 205 data kemudian dibagi menjadi 2 data yaitu data training sejumlah 164 data dan data testing sejumlah 41 data dengan perbandingan 80% dan 20% diproses dengan menggunakan algoritma C4.5 dengan akurasi

- yang didapat sebesar 78,75%.
- b. Prediksi Kebangkrutan Perusahaan Menggunakan Algoritma C4.5 Berbasis Forward Selection (Saleh, 2017). Penelitian ini membahas kebangkrutan perusahaan dengan menggunakan eksperimen data diperoleh dari dataset global yang datanya bervariasi jumlah recordnya yang berjumlah 250 record dan 150 record diproses dengan menggunakan algoritma C4.5 dengan akurasi yang didapat sebesar 99.60%.
- c. Penerapan Algoritma C4.5 Untuk Prediksi Tingkat Kompetensi Karyawan PT Multi Strada Arah Sarana (Wahyono dkk, 2018). Penelitian ini membahas kompetensi karyawan dengan menggunakan eksperimen data diperoleh dari hasil tes kompetensi karyawan PT Multistrada Arah Sarana yang datanya dikumpulkan berjumlah 257 data karyawan kemudian dibagi menjadi 2 data yaitu data training sejumlah 206 data dan data testing sejumlah 51 data dengan perbandingan 80% dan 20% diproses dengan menggunakan algoritma C4.5 dengan akurasi yang didapat sebesar 80,39%.
- d. Penerapan Algoritma Klasifikasi Data Mining C4.5 Pada Dataset Cuaca Wilayah Bekasi (Novandya dkk, 2017). Penelitian ini membahas cuaca wilayah beksi dengan menggunakan eksperimen data diperoleh dari situs World Weather Online diproses dengan menggunakan algoritma C4.5 dengan akurasi yang didapat sebesar 88,89%.

### 3. Kerangka Studi

Adapun kerangka pemikiran dari penelitian yang dilakukan adalah sebagai berikut:



### 4. Desain Penelitian/Metodologi

Jenis penelitian ini menggunakan model eksperimen, merupakan menggunakan data dalam penelitian dan menghasilkan kesimpulan yang mampu dibuktikan oleh pengamatan atau percobaan (Sunge, 2018).

Penelitian ini menggunakan data primer yang diperoleh secara langsung di PT. Shuangying International Indonesia karena peneliti bekerja sebagai karyawan di PT. Shuangying International Indonesia. Setelah data dikumpulkan berjumlah 112

data kemudian dibagi menjadi 2 yaitu data training sejumlah 90 data dan data testing sejumlah 22 data dengan perbandingan 80% dan 20%. Hasil data training digunakan untuk memperoleh hasil klasifikasi untuk prediksi pengendalian kualitas dalam bentuk pohon keputusan **dan data testing digunakan tingkat akurasi dari hasil klasifikasi tersebut.**

### 5. Hasil Penelitian dan Pengujian

Dari proses klasifikasi dengan menggunakan *Knowledge Discovery in Databases* (KDD) kemudian menganalisa produk cacat.

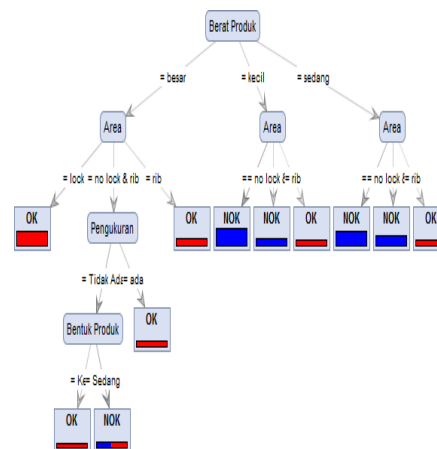
**Tabel 5.1** Kategori Produk Cacat pada PT. Shuangying International Indonesia.

| No | Atribut       | Penentuan Calon Cabang |                      |                    |
|----|---------------|------------------------|----------------------|--------------------|
|    |               | Berat Besar            | Berat Sedang         | Berat Kecil        |
| 1  | Berat Produk  | Berat Besar            | Berat Sedang         | Berat Kecil        |
| 2  | Bentuk Produk | Bentuk Besar           | Bentuk Sedang        | Bentuk Kecil       |
| 3  | Pengukuran    | Pengukuran ada         | Pengukuran tidak ada |                    |
| 4  | Ukuran Cacat  | Ukuran Besar           | Ukuran Kecil         | Ukuran tidak ada   |
| 5  | Posisi        | Posisi Luar            | Posisi Dalam         |                    |
| 6  | Area          | Area Lock              | Area rib             | Area no lock & rib |

Dari hasil yang ada kemudian dikategorikan dengan *variable, atribut*

Dan dilanjutkan penelitian untuk dilakukan proses validasi menggunakan fitur *cross validation*. Berikut merupakan permodelan yang ada didalam *cross validation* :

Pada permodelan *cross validation* didalamnya terdapat dua bagian, yaitu bagian *training* (digunakan untuk algoritma klasifikasi *Decision Tree*) dan *testing* (menggunakan fitur *Apply Model* untuk mengaplikasikan model pada data *testing* dan fitur *Performance* untuk menampilkan *confusion table*, yang digunakan untuk menampilkan hasil dari *accuracy, recall, precision*, dan nilai AUC.



Maka terbentuk simpul-simpul diperoleh *decision tree* untuk klasifikasi prediksi produk cacat pada gambar 5.3

## 6. Kesimpulan

Kesimpulan dari penelitian ini adalah hasil analisa data mining dari produk cacat pada PT. Shuangying International Indonesia dengan metode klasifikasi (C4.5).

Data mining dapat diaplikasikan sebagai pengetahuan, informasi untuk mengatasi masalah yang telah teridentifikasi pada penelitian ini. Dari hasil data mining produk cacat, dapat terurai menjadi pengetahuan dan prediksi sehingga hasil dari pengolahan data dapat menjadi sumber informasi dalam produk cacat. Dari 90 data training diperoleh menjadi produk OK (37) dan NOK (57). Berdasarkan evaluasi dan pengujian prediksi yang telah dilakukan, analisa data mining dengan metode klasifikasi untuk produk cacat pada PT. Shuangying International Indonesia, menggunakan Algoritma C4.5 dengan Confusion Matrix menghasilkan accuracy sebesar 95,56%, recall sebesar 95,00%, dan precision sebesar 95,00%. Hasil yang didapat dari pengolahan kurva ROC sebesar 1.000 menggunakan data training dengan tingkat akurasi Excellent Clasification.

## Daftar Pustaka

- Aprilla C, Dennis., Baskoro, Donny Aji., Ambarwati, Lia., Wicaksana, I Wayan Simri. (2013). Belajar Data Mining Dengan RapidMiner, 5(4), 1–5. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Cahyadi, D. (2014). Analisis Parameter Operasi Pada Proses Plastik Injection Molding Untuk Pengendalian Cacat Produk. *Jurnal Mesin Teknologi*, 8(2), 8–16. Retrieved from <https://jurnal.umj.ac.id/index.php/sintek/article/view/161/143>
- Dai, Wei., Ji, W. (2014). A Map Reduce Implementation of C4.5 Decision Tree Algorithm, 7(1), 49–60.
- Dewi, Shanty Kusuma. (2012). Minimasi Defect Produk Dengan Konsep Six Sigma. *Jurnal Teknik Industri*, 13, 43–50.
- Gullo, Francesco. (2015). From Patterns in Data to Knowledge Discovery: What Data Mining Can Do. *Physics Procedia*, 62, 18–22. <https://doi.org/10.1016/j.phpro.2015.02.005>
- Han, Jiawei., Kamber, Micheline ., Pei, Jian. (2012). *Data Mining: Concepts and Techniques*. San Francisco, CA, itd: Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-381479-1.00001-0>
- Kotu, Vijay., Deshpanda, Bala. (2015). Predictive Analytics and Data Mining. In *Predictive Analytics and Data Mining* (pp. 1–16). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-801460-8.00001-X>
- Mardi, Yuli. (2017). Data Mining: Klasifikasi Menggunakan Algoritma C4.5. *Jurnal Edik Informatika*, 2(2), 213–219.
- Novandya, Adhika., Oktria, Isni. (2017). Penerapan Algoritma Klasifikasi Data Mining C4.5 Pada Dataset Cuaca Wilayah Bekasi. *Jurnal Format*, 6(2), 98–106.
- Raharja, Yosoa Putra. (2014). Rancang Bangun Sistem Rekomendasi Beasiswa Menggunakan Algoritma Klasifikasi C4.5 pada Universitas Dian Nuswantoro. *Undinus*, 1–4. Retrieved from <http://eprints.dinus.ac.id/13408/>
- Rohman, Abdul. (2016). Komporasi Metode Klasifikasi Data Mining Untuk Prediksi Penyakit Jantung. *Neo Teknika*, 2(2), 21–28.
- Saifudin, Aries., Wahono, Romi Satrio. (2015). Penerapan Teknik Ensemble untuk Menangani Ketidakseimbangan Kelas pada Prediksi Cacat Software. *Journal of Software Engineering*, 1(1), 28–37. [https://doi.org/10.1016/S1896-1126\(14\)00030-3](https://doi.org/10.1016/S1896-1126(14)00030-3)
- Saleh, Hamsir. (2017). Prediksi Kebangrutan Perusahaan Menggunakan Algoritma C4.5 Berbasis Forward Selection, 9, 173–180.
- Silalahi, Kristiani Desri., Murfi, Hendri., Satria, Yudi. (2017). Studi Perbandingan Pemilihan Fitur untuk Support Vector Machine pada Klasifikasi Penilaian Risiko Kredit, 1(2), 119–136.
- Sinambela, Mutiara Shany., Mayadewi, RA Paramita., Rosely, Ely. (2016). Sistem Pendukung Keputusan Peminatan Jurusan Siswa Di SMA Menggunakan Algoritma Klasifikasi Data Mining C4.5, 2(3), 858–866.
- Singh, Sonia., Gupta, Prinyanka. (2014). Comparative Study ID3 , CART And C4.5 Decision Tree Algorithm : A Survey, 27(27), 97–103.