

**VALIDASI SILANG PENGELOMPOKAN JENIS KALENG  
MENGUNAKAN METODE *DECISION TREE***

**SKRIPSI**

**Sebagai Salah Satu Syarat Untuk Memperoleh Gelar  
Sarjana Sains Bidang Studi Matematika**



**Oleh:**

**MEGAWATI  
NIM. 08011181621078**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS SRIWIJAYA  
2021**

**LEMBAR PENGESAHAN**

**VALIDASI SILANG PENGELOMPOKAN JENIS KALENG  
MENGUNAKAN METODE *DECISION TREE***

**SKRIPSI**

**Sebagai Salah Satu Syarat Untuk Memperoleh Gelar  
Sarjana Sains Bidang Studi Matematika**

**Oleh :**

**MEGAWATI  
NIM. 08011181621078**

**Pembimbing Kedua**

**Indralaya, Januari 2021  
Pembimbing Utama**



**Des Alwine Zavanti, M.Si  
NIP.197012041993022001**



**Dr. Yulia Resti, M.Si  
NIP.197307191997022001**

**Mengetahui,  
Ketua Jurusan Matematika**



**Drs. Sugandi Yahdin, M.M  
NIP.195807271986031003**

## LEMBAR PERSEMBAHAN

### MOTTO

**"SESUNGGUHNYA SESUDAH KESULITAN ITU ADA KEMUDAHAN"  
(Q.S AL-INSYIRAH : 5-6)**

**"JANGAN PERGI MENGIKUTI KEMANA JALAN AKAN BERJUNG.  
BUAT JALANMU SENDIRI DAN TINGGALKANLAH JEJAK"**

**"BARANG SIAPA YANG TIDAK MAMPU MENAHAN LELAHNYA  
BELAJAR, MAKA IA HARUS MAMPU MENAHAN PERIHNYA  
KEBODOHAN"  
(HR. IMAM AS-SYAFI' I)**

*Skripsi ini kupersembahkan untuk :*

- ALLAH SWT
- Rasulullah Muhammad SAW
- Kedua Orangtuaku
- Keluarga Besarku
- Semua Guru dan Dosenku
- Sahabat-sahabatku
- Almamaterku

## KATA PENGANTAR

**Assalammu'alaikum wr. wb**

Puji dan syukur penulis panjatkan kepada Allah SWT atas berkat, rahmat dan karunia-Nya yang luar biasa sehingga penulis bisa menyelesaikan skripsi yang berjudul “**Validasi Silang Pengelompokkan Jenis Kaleng Menggunakan Metode *Decision Tree***” dapat berjalan dengan baik dan selesai pada waktunya. Shalawat serta salam semoga senantiasa selalu tercurahkan kepada Nabi Muhammad *Shallallahu 'alaihi wa sallam* yang telah mengantarkan manusia dari zaman kegelapan ke zaman yang terang benderang. Skripsi ini sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains bidang studi Matematika di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya. Dengan segala hormat, kasih sayang dan kerendahan hati penulis mengucapkan terima kasih yang sedalam-dalamnya kepada Kedua Orang Tua, yaitu Ayah, **Kamalludin** dan Ibu, **Lindawati** dengan segenap cinta, kasih sayang berlimpah, nasehat, didikan dan dukungan moral maupun spiritual yang tak pernah henti untuk keberhasilan penulis. Selain itu penulisan skripsi ini tidak lepas dari bantuan berbagai pihak, untuk itu penulis menyampaikan terima kasih kepada:

1. Bapak **Drs. Sugandi Yahdin, M.Si**, selaku Ketua Jurusan Matematika dan selaku dosen yang sangat memperhatikan penulis dan rekan-rekan lain dengan bimbingan, arahan, kritik dan saran yang telah diberikan selama penulis belajar di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.

2. Ibu **Des Alwine Zayanti, M.Si**, selaku Sekretaris Jurusan Matematika FMIPA Universitas Sriwijaya sekaligus Dosen Pembimbing Kedua yang bersedia meluangkan waktu dan memberikan arahan, bimbingan, dorongan, motivasi, kritik dan masukan selama penulis dalam masa kuliah terutama dalam proses pengerjaan skripsi.
3. Ibu **Dr. Yulia Resti, M.Si**, selaku Pembimbing Utama yang telah bersedia meluangkan banyak waktu, pikiran, tenaga, dan motivasi sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
4. Bapak **Drs. Ngudiantoro, M.Si**, selaku Dosen Pembimbing akademik yang telah memberikan bimbingan selama penulis berstatus sebagai mahasiswa di Jurusan Matematika FMIPA Universitas Sriwijaya.
5. Ibu **Novi Rustiana Dewi, M.Si**, Ibu **Eka Susanti, M.Si** dan Ibu **Endang Sri Kresnawati, M.Si** selaku Dosen Penguji yang telah bersedia meluangkan waktu dalam memberikan tanggapan, kritik dan saran yang bermanfaat dalam perbaikan dan penyelesaian skripsi ini.
6. **Seluruh Dosen di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam** yang telah memberikan ilmu yang bermanfaat, nasehat serta bimbingan selama penulis menjalani perkuliahan.
7. **Pak Irwansyah** dan **Ibu Hamidah** Matematika selaku admin jurusan yang telah membantu penulis selama masa perkuliahan.
8. **Ayukku Lia Andriani dan adik-adikku Yulna Lesti, Oktalika, M.Wahyu Kodar dan M.Zayn Aldito Akbar** yang telah memberikan semangat, kasih

dan sayangnya dalam mengerjakan skripsi serta motivasi yang tiada henti diberikan.

9. **Tim skripsi terbaik, anbil, eling, nana, rendi, eko, irvan, ogi, jekta, ari, tiwul, muti, fero, anggun, indah, widya, riska, hariani, Seluruh angkatan 2016 serta teman-teman SD, SMP seperjuangan** yang telah memberikan semangat, motivasi, do'a, dan kerja sama yang baik dalam penyusunan skripsi ini.
10. **Sahabat terbaikku, Twelve Family, Chili genk, 7-icon dan jannah project** yang selalu menyemangatiku dan memberikan dukungan.
11. **Kakak-kakak tingkat** angkatan 2012, 2013,2014 dan 2015 serta **adik-adik tingkat** semua angkatan 2017, 2018 dan 2019.
12. Keluargaku di organisasi terbaik, **BPH LDF KOSMIC 2017/2018 dan MUDO SOSIAL EKPEDISI** terimakasih atas pengalaman selama penulis masih studi.
13. Semua pihak yang telah membantu penulis dalam menyelesaikan skripsi yang tidak dapat disebutkan satu persatu. Terimakasih atas semua dukungan, Penulis mengharapkan kritik dan saran untuk meningkatkan kualitas dari skripsi ini dan semoga dapat bermanfaat bagi semua yang membacanya.

Wassalammu'alaikum wr. wb

Indralaya, Januari 2020

Penulis

**CROSS VALIDATION IN GROUPING TYPES OF CANS USING THE  
*DECISION TREE* METHOD**

**By:**

**MEGAWATI  
08011181621078**

**ABSTRACT**

Cans are storage places for food and non-food items made of aluminum. Canned waste is difficult to decompose on the ground so that a recycling process needs to be carried out so that it does not cause problems for the environment. The process of grouping the cans is the initial process in recycling. This study classifies cans based on RGB images using the *Decision Tree*. The decision tree explains the effect of attributes by depicting trees. The data consisted of 250 samples of cans given treatment factors in the form of one lighting angle, one type of light lamp, and two conveyor belt speeds. Delivery of treatments on the data is conveyor belt speed 1, lamp1, angle 30°, and conveyor belt speed 2, lamp 1, angle 30°. With cross validation 5 times get 1 fold as data testing and training 4 times data tested over and over. From the results of cross validation using the *Decision Tree* the greatest accuracy rate is 75.25%.

**Keywords:** cans, decision tree, cross validation

## VALIDASI SILANG PENGELOMPOKKAN JENIS KALENG MENGUNAKAN METODE *DECISION TREE*

Oleh :

MEGAWATI  
08011181621078

### ABSTRAK

Kaleng merupakan tempat penyimpanan bahan pangan dan non-pangan yang terbuat dari bahan alumunium. Limbah kaleng sulit terurai di tanah sehingga perlu dilakukan proses daur ulang agar tidak menimbulkan masalah bagi lingkungan. Proses pengelompokkan kaleng menjadi proses awal dalam daur ulang. Penelitian ini mengklasifikasikan kaleng berdasarkan citra RGB menggunakan metode *Decision Tree*. *Decision Tree* menjelaskan pengaruh atribut dengan penggambaran pohon. Data terdiri dari 250 sampel kaleng yang diberikan faktor perlakuan berupa satu sudut pencahayaan, satu jenis cahaya lampu, dan dua kecepatan *Conveyor Belt*. sehingga didapatkan kombinasi perlakuan pada data adalah kecepatan *conveyor belt* 1, lampu1, sudut  $30^0$ , dan kecepatan *conveyor belt* 2, lampu 1, sudut  $30^0$ . Dengan *5-fold cross validation* didapatkan 1 *fold* sebagai data *testing* dan 4 *fold* data *training* yang diuji secara berulang. Dari hasil *cross validation* dengan menggunakan *Decision Tree* tingkat akurasi terbesar 75.25%.

**Kata kunci** : kaleng, decision tree, validasi silang



## DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	<b>i</b>
<b>LEMBAR PENGESAHAN</b> .....	<b>ii</b>
<b>LEMBAR PERSEMBAHAN</b> .....	<b>iii</b>
<b>KATA PENGANTAR</b> .....	<b>iv</b>
<b>ABSTRACT</b> .....	<b>vii</b>
<b>ABSTRAK</b> .....	<b>viii</b>
<b>DAFTAR ISI</b> .....	<b>ix</b>
<b>DAFTAR TABEL</b> .....	<b>xii</b>
<b>DAFTAR GAMBAR</b> .....	<b>xiv</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xv</b>
<b>BAB I PENDAHULUAN</b>	
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	3
1.3. Pembatasan Masalah .....	3
1.4. Tujuan Penelitian .....	4
1.5. Manfaat Penelitian .....	4
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1. Kemasan Kaleng .....	5
2.1.1. Kaleng Timah ( <i>tin-plate</i> ) .....	5
2.1.2. Kaleng Aerosol .....	6
2.1.3. Kaleng Aluminium.....	6

2.2. Pengolahan Citra Digital .....	6
2.3. Citra RGB .....	7
2.4. <i>K-Fold Cross Validation</i> .....	7
2.5. <i>Decision Tree</i> .....	8
2.5.1. Algoritma ID3 .....	9
2.5.2. Algoritma C4.5 .....	10
2.6. Probabilitas .....	12
2.7. Diskritisasi Data .....	14
2.8. Tingkat Akurasi Metode .....	14

### **BAB III METODOLOGI PENELITIAN**

3.1. Tempat .....	16
3.2. Waktu .....	16
3.3. Metode Penelitian .....	16

### **BAB IV HASIL DAN PEMBAHASAN**

4.1. Mengelompokkan Data Dengan <i>Cross Validation</i> .....	21
4.2. Diskritisasi Data .....	21
4.3. <i>Decision Tree C4</i> .....	24
4.4. Mengelompokkan Data <i>Berhirarki</i> Variabel .....	28
4.5. Diskritisasi Data .....	29
4.6. <i>Decision Tree C4.5 Berhirarki</i> .....	30

### **BAB V KESIMPULAN DAN SARAN**

5.1. Kesimpulan .....	36
5.2. Saran .....	36

<b>DAFTAR PUSTAKA .....</b>	<b>37</b>
<b>LAMPIRAN .....</b>	<b>38</b>

## DAFTAR TABEL

	Halaman
Tabel 2.1	Skenario Uji Akurasi Dengan Metode <i>5-Fold cross validation</i> .....8
Tabel 2.2	Tingkat Akurasi Metode.....14
Tabel 4.1	Data Keseluruhan Nilai RGB kaleng .....21
Tabel 4.2	Data Komposisi <i>Fold 1</i> .....21
Tabel 4.3	Range Nilai RGB <i>Fold 1</i> .....22
Tabel 4.4	Data Diskritisasi Komposisi <i>Fold 1</i> .....23
Tabel 4.5	Data Nilai Entropy <i>Fold 1</i> .....25
Tabel 4.6	Data Peluang Tiap Atribut.....26
Tabel 4.7	Klasifikasi Metode <i>Decision Tree C4.5</i> .....28
Tabel 4.8	Tingkat Akurasi <i>Decision Tree C4.5</i> .....28
Tabel 4.9	Data komposisi <i>Fold 1 Berhirarki</i> .....29
Tabel 4.10	Range nilai RGB <i>Fold 1 Berhirarki</i> .....29
Tabel 4.11	Data Diskritisasi Komposisi <i>Fold 1 Berhirarki</i> .....30
Tabel 4.12	Data Nilai Entropy <i>Fold 1 Berhiarki</i> .....32
Tabel 4.13	Data Peluang Tiap Atribut.....32
Tabel 4.14	Klasifikasi Metode <i>Decision Tree C4.5</i> . .....33
Tabel 4.15	Tingkat Akurasi <i>Decision Tree C4.5</i> .....33
Tabel 4.16	Tingkat akurasi <i>Decision Tree C4.5</i> .....34
Tabel 4.17	Rata-Rata Tingkat Akurasi .....35

## DAFTAR GAMBAR

	Halaman
Gambar 4.1 Pohon Keputusan Node .....	27
Gambar 4.2 Pohon Keputusan Node 1 <i>Berhirarki</i> .....	33

## DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Data Fold 2.....	40
Lampiran 2 Range Data <i>Fold 2</i> .....	40
Lampiran 3 Range Himpunan RGB Data <i>Fold 2</i> .....	40
Lampiran 4 Data Diskritisasi <i>Fold 2</i> .....	41
Lampiran 5 Nilai Entropy <i>Fold 2</i> .....	41
Lampiran 6 Nilai <i>Gain</i> Komposisi <i>Fold 2</i> .....	42
Lampiran 7 Pohon Keputusan <i>Fold 2</i> .....	42
Lampiran 8 Hasil Klasifikasi Keputusan <i>Fold 2</i> .....	43
Lampiran 9 Akurasi <i>Decision Tree</i> .....	43
Lampiran 10 Rata-Rata Akurasi <i>Decision Tree</i> .....	44

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Semakin berkembangnya perindustrian di era sekarang ini membuat proses pengemasan barang menjadi hal yang sangat dibutuhkan dari segi pengawetan dan kepraktisan. Tetapi dampak terhadap produk di dalamnya dan lingkungan juga harus diperhatikan. Kaleng adalah salah satu tempat penyimpanan makanan, minuman dan bahan *non*-pangan lainnya yang terbentuk dari aluminium atau logam, dan merupakan salah satu bahan pengemasan yang paling banyak digunakan dalam kehidupan sehari-hari (Cenadi,2000).

Dalam proses pengolahan limbah terdapat 3 cara yaitu *reuse* (pemakaian kembali), *reduce* (pengurangan pemakaian) dan *recycle* (daur ulang). Limbah sampah kaleng menjadi limbah yang sulit terurai di tanah, sehingga perlu dilakukan proses *recycle* (daur ulang). Dalam proses daur ulang kaleng dikelompokkan berdasarkan jenisnya yaitu kaleng *tin-plate* untuk makanan, kaleng aluminium untuk minuman, dan kaleng aerosol untuk cat, parfum, serta pengharum ruangan. Pengelompokan tersebut bisa didasari oleh teknik citra digital. Proses teknik citra digital merupakan teknik dimana informasi didapat melalui gambar dari hasil ekstrasi gambar menjadi nilai *Red Green Blue* (RGB). Dalam pengklasifikasian objek terdapat beberapa metode yang dapat dilakukan, pada penelitian ini dilakukan uji dengan metode yaitu *Decision Tree* (DT). Konsep dasar algoritma DT adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan.

Validasi Silang adalah cara untuk menguji suatu model, salah satu proses dalam validasi silang adalah *K-fold cross validation*. *K-fold cross validation* merupakan salah satu metode dimana data dibagi  $n$  – data dan dilakukan pengujian secara silang dan acak yang dilakukan berulang sehingga dapat diketahui rata-rata keberhasilan suatu data. Menurut Sofia,dkk (2018) yang mengimplementasikan *Fuzzy Decision Tree* untuk memprediksikan gagal ginjal kronis dengan *K-fold Cross Validation* didapatkan akurasi tertinggi adalah 98,28% dengan (K=6) dan 6 atribut yang memiliki pengaruh terbesar dari 11 atribut yang digunakan.

Menurut Pangestika (2019) bahwa penerapan metode *Fuzzy Naive Bayes* dalam mengelompokkan jenis kaleng berdasarkan RGB dengan komposisi perbandingan data *training* dan data *testing* digunakan 3 komposisi yakni, 30% : 70%, 50% : 50%, dan 70% : 30% didapatkan tingkat akurasi tertinggi sebesar 64% pada komposisi pertama yaitu 30% data *training* dan 70% data *testing*. Menurut Sari (2020) yang menerapkan validasi silang dalam mengklasifikasikan jenis kaleng dengan citra RGB menggunakan Metode *Fuzzy Naive Bayes* dengan *K-Fold cross Validatio* (K=10) didapatkan akurasi tertinggi adalah 74.64% pada data kecepatan *conveyor belt 1*, lampu1, sudut  $90^0$ .

Dari penelitian sebelumnya hasil dengan melakukan perbandingan komposisi memiliki akurasi yang masih kecil daripada menggunakan *K-Fold cross Validation*. Sehingga pada penelitian kali ini akan dilakukan perbandingan tingkat akurasi DT dengan *K-Fold cross validation*.



## 1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah :

1. Bagaimana pengelompokan jenis kaleng berdasarkan citra *RGB* dengan metode *K-fold cross validation* terhadap metode DT?
2. Bagaimana hasil pengelompokkan jenis kaleng berdasarkan citra *RGB* dengan metode DT?

## 1.3 Pembatasan Masalah

Pembatasan masalah pada penelitian ini adalah:

1. Sampel yang digunakan dibatasi sebanyak 250 kaleng yang terdiri dari 74 kaleng *tin plate*, 83 kaleng aluminium, dan 93 kaleng aerosol dengan posisi foto tiap kaleng yaitu depan, belakang dan samping.
2. Faktor-faktor yang mempengaruhi hasil foto pada penelitian ini adalah kecepatan *Conveyor Belt*, pencahayaan lampu dan sudut pencahayaan lampu.
3. kecepatan *Conveyor Belt* terbagi 2 yaitu kecepatan 1 sebesar 0.181m/s dan kecepatan 2 sebesar 0.086 m/s.
4. Sudut yang di gunakan dalam pengambilan foto kaleng adalah  $30^0$ .

#### **1.4 Tujuan Penelitian**

1. Mengaplikasikan metode DT untuk mengelompokkan jenis kaleng berdasarkan citra RGB menggunakan metode validasi silang *K-Fold*.
2. Mengetahui perbandingan tingkat akurasi tertinggi antara metode DT.

#### **1.5 Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan manfaat, yaitu :

1. Bagi Jurusan Matematika  
Memotivasi mahasiswa lainnya untuk melakukan penelitian menggunakan metode pengelompokkan lainnya.
2. Bagi Peneliti lain  
Sebagai bahan referensi mengenai penerapan tentan mengelompokkan suatu objek berdasarkan citra RGB dengan metode DT.

## DAFTAR PUSTAKA

- Arifin, Z. 2016. *Klasifikasi Kematangan Buah Jambu Biji Merah Menggunakan Naive Bayes Berdasarkan Warna RGB* [skripsi]. Semarang: Fakultas Ilmu Komputer. Universitas Dian Nuswantoro, Semarang.
- Cenadi, C. S. (2000). Peranan Desain Kemasan dalam Dunian Pemasaran. *Nirwana*, 92-103.
- Defiyanti, S., & D. L. Crispina Pardede. (2008). Perbandingan kinerja algoritma id3 dan c4.5 dalam klasifikasi spam-mail. *ReCALL*.
- Haryati., Sudarsono., dan Suryana. (2015). Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5, 11(2).
- Hogg, R. V., and Tannis, E.A. 2006. *Probability And Statistical Inference*. Ed ke-7. United States of America : Pearson Education, Inc.
- Kandel, A., Pacheco, R., Martina, A., & Khator, S.. (1996). *The Foundations of Rule-based Computations in Fuzzy Models Pedryez (Ed). Fuzzy Modelling Paradigme and Practice (pp. 231-263). United State; Kluwer Academic Publishers.*
- Kusmanto, R., dan Tompunu, A. N. 2011. Pengolahan Citra Digital Untuk Mendeteksi Obyek Menggunakan Pengolahan Warna Model Normalisasi RGB. Di dalam : *Seminar Nasional Teknologi Informasi & Komunikasi Terapan 2011*.
- Kusumadewi, Sri. 2003. *Analisis & Desain Sistem Fuzzy Menggunakan Toolbox Matlab*. Yogyakarta: Graha Ilmu.
- Larose, D. T., 2005, *Discovering Knowledge in Data: An Introduction to Data Mining*, John Wiley & Sons Inc.
- Liang, G. (2005). *A comparative study of three Decision Tree algorithms: ID3, Fuzzy ID3 and Probabilistic Fuzzy ID3*. Netherlands: Erasmus University.
- Manik, F.Y., dan Saragih, K. S. (2017). *Klasifikasi Belimbing Menggunakan Naive Bayes Berdasarkan Fitur Warna RGB*. *IJCSS*, Vol.11 No.1 : 99-108.
- Professional, A. P. (1994). *Fuzzy Systems Handbook* (Harcourt Brace & Company (ed.); BYTE Magaz). Academic Press Limited.

- Putra Dama., Purnaman., dan Putri Singgih., 2018, *Diagnosa Penyakit Mata Dengan Fuzzy Logic dan Naive Bayes*, Vol 6, No 1.
- Putri , E. (2012, 12 November 2018). Retrieved from <https://ekaputri12.wordpress.com/2012/12/19/aerosol-kaleng/>
- Rahimah S. 2011. *Kemasan Kaleng*. <http://blogs.unpad.ac.id/souvia/files/2011/03/KEMASAN-LOGAM-2011.pdf> [08 September 2018]
- Rahmawati, D. (2013, 2 November 2018). Retrieved from <http://rahmawati.blogspot.com/2013/07/kemasan-kaleng-jenis-kaleng.html>.
- Riza Khadafy, A., & Satria Wahono, R. (2015). *Penerapan Naive Bayes untuk Mengurangi Data Noise pada Klasifikasi Multi Kelas dengan Decision Tree*. *Journal of Intelligent Systems*, 1(2). <http://journal.ilmukomputer.org>.
- Supriyanti, W., Kusriani, & Amborowati, A. (2016). *Perbandingan Kinerja Algoritma c4.5 Dan Naive Bayes Untuk Ketepatan Pemilihan Konsentrasi Mahasiswa*. *Jurnal INFORMA Politeknik Indonusa*, 1(3), 61–67.
- Taufiq, Ghofar. (2016). Implementasi Logika *Fuzzy* Tahani untuk Model Sistem Pendukung Keputusan Evaluasi Kinerja Karyawan. *Jurnal Pilar Nusa Mandiri*, Vol.XII No.1:13-14.
- Walpole, Ronald E. (1995). *Pengantar Statistika*. Ed ke-3. Jakarta: Penerbit PT. Gramedia Pustaka Utama.
- Wang, Li-Xin. 1997. *A course in Fuzzy Systems and Control*. London: Prentice-Hall International.