

## Penerapan Algoritma *Decision Tree* dan *Feature Selection Forward* untuk Analisis Kepuasan Pemasang Iklan Terhadap Pelayanan Iklan

Jimi Ali Baba, Guna Yanti Kemala Sari Siregar Pahu, Ricci Herdian Saputra  
Jurusan Sistem Informasi, STMIK Pringsewu

Correspondence Author: [jimialibaba@gmail.com](mailto:jimialibaba@gmail.com)

### Abstrak

Media cetak adalah media informasi berupa koran, majalah, dan lain sebagainya dimana media cetak tersebut di pergunakan oleh perorang atau perusahaan untuk mencari informasi baik berupa berita maupun iklan yang di sajikan. Iklan sebagai media informasi sangat penting untuk di pergunakan dalam berpromosi baik perusahaan maupun perorangan, namun terkadang pelayanan terhadap pemasang iklan kurang merasa terlayani dengan baik oleh sebagian perusahaan yang bergerak di industri cetak. Dataset kepuasan konsumen di dapat dari data primer yaitu berupa data kuisioner yang di sajikan sehingga terbentuk menjadi dataset yang di gunakan untuk menguji kepuasan pelayanan iklan. Oleh karena perlu di terapkan suatu metode untuk menganalisis seberapa puaskan konsumen terhadap pelayanan yang di lakukan oleh pihak media cetak, dalam hal ini di lakukan penerapan metode analisis dengan menggunakan Algoritma C4.5 dan Feature Selection Forward untuk mengukur tingkat kepuasan pelayanan iklan terhadap konsumennya. Penerapan metode C4.5 dan Feature Selection Forward sebagai peningkatan akurasi sangat baik di lakukan dalam menganalisis kepuasan pelayanan tersebut, terbukti dengan menggunakan algoritma C4.5 84.50% kepuasan pelayanan iklan yang tersaji, sedangkan algoritma C4.5 tersebut di reduksi dengan Feature Selection Forward dapan meningkatkan akurasi mencapai 90% dan termasuk kedalam Excelent Classification.

Keyword: *Decision tree*, C4.5, *feature selection forward*, media cetak, data mining

### 1. PENDAHULUAN

Lampung Post di bawah manajemen PT. Masa Kini Mandiri adalah perusahaan bergerak di bidang perdagangan industri *pers* dan percetakan. Lampung Post berdiri sejak 10 Agustus 1974. Pada awal berdirinya PT. Masa Kini Mandiri beroperasi dengan karyawan sebanyak 50 orang, dan perkembangan saat ini dapat beroperasi dengan jumlah seluruh karyawan sebanyak 208 orang melalui jalur bisnis utamanya adalah harian umum di Provinsi Lampung. Kantor pusat PT. Masa Kini Mandiri terletak di Jalan Soekarno Hatta No. 108, Rajabasa, Bandar Lampung sedangkan kantor perwakilannya terletak di Gedung Media Indonesia.

Sebagai media beriklan, media massa cetak pernah mengalami masa kejayaan, ketika beriklan di televisi dilarang. Seiring dengan deregulasi iklan televisi, media cetak menjadi media beriklan yang kalah favorit dibandingkan media televisi, walaupun media cetak mempunyai spesifikasi yang berbeda dengan media lainnya.

Lampung Post yang mempunyai pelanggan tetap pemasang iklan harus bekerja ekstra untuk memberikan pelayanan yang maksimal supaya pelanggan-pelanggan iklan tidak berpaling ke lain media cetak, data yang di dapat dari penelitian ini yang berlangganan memasang iklan sekitar 40301 pelanggan dari tahun 2008 hingga sekarang dan seiring berkembangnya industri advertaising seakan pelanggan iklan tergerus oleh kemajuan teknologi di mana pelanggan sudah bisa memasang iklan gratis di layanan *online*. Untuk saat ini pelanggan iklan Lampung Post hanya mencapai 50 sampai 200 pelanggan untuk tiap minggunya, dimana tiap-tiap pelanggan tersebar seluruhnya di lampung dan terdapat juga dari pulau jawa sebagai pemasang iklan tersebut. Dari permasalahan-permasalahan tersebut divisi iklan harus bekerja ekstra supaya pelanggan iklan tidak berpaling salah satunya yaitu memberikan pelayanan semaksimal mungkin.

Semakin berkembangnya daerah Lampung mengakibatkan tumbuhnya media-media cetak yang notabennya mengharapakan pemasukan dari iklan, pendapatan iklan yang sangat menggiurkan bagi kalangan media cetak sehingga pesaing-pesaing di dunia industri cetak berlomba memberikan pelayanan ekstra terhadap konsumen pemasang iklan tersebut.

Menurut sebagian pihak, pelayanan iklan pada PT. Masa Kini Mandiri (Lampung Post) sungguh kurang dari yang diharapkan. Pihak tersebut menilai pelayanan yang diberikan kurang maksimal. Permasalahan lainnya yang juga sering dijumpai yaitu terkait keterlambatan terbitnya iklan, keterlambatan waktu terbitnya iklan juga di soroti oleh sebagian pihak atas ketidak profesional organisasi tersebut.

Menyikapi anggapan tersebut, maka penulis merasa perlu untuk melakukan analisis kepuasan konsumen terhadap pelayanan iklan yang ada.

## 2. METODE PENELITIAN

### 1. Pengumpulan data

Pada tahap ini ditentukan data yang akan diproses. Mencari data yang tersedia, memperoleh data tambahan yang dibutuhkan, mengintegrasikan semua data kedalam data set, termasuk variabel yang diperlukan dalam proses.

### 2. Pengolahan data awal

Pada tahap ini dilakukan penyeleksian data, data dibersihkan dan ditransformasikan kebentuk yang diinginkan sehingga dapat dilakukan persiapan pembuatan model.

### 3. Metode yang diusulkan

Pada mana yang berhubungan dengan satu sama lainnya. Setelah data dianalisis lalu diterapkan model-model yang sesuai dengan jenis data.

### 4. Eksperimen dan pengujian metode

Pada tahap ini model yang diusulkan akan diuji untuk melihat hasil berupa *rule* yang akan dimanfaatkan dalam pengambilan keputusan.

### 5. Evaluasi dan validasi

Pada tahap ini dilakukan evaluasi terhadap model yang diterapkan untuk mengetahui tingkat keakurasian model. Berikut gambar langkah-langkah dalam tahap penelitian dimana pada tahap ini data dianalisis dan dikelompokkan berdasarkan variable.



Gambar 1. Tahapan penelitian

## 3. HASIL DAN ANALISA

Data uji yang digunakan dalam penelitian ini bersumber pada dataset privat dimana data diperoleh dari hasil kuisioner yang disebarkan kepada pelanggan iklan sebanyak 200 respon atau pelanggan iklan yang berkunjung ke Lampung Post. Dari koleksi dataset tersebut terdapat 12 Variabel/Attribut dimana variable yang menjadi *root note* utama yang paling mewakili kepuasa terhadap pelanggan iklan pada Lampung Post adalah pada Variabel Tanggibles. Terdapat 84.50% akurasi untuk lima aspek dan di tambah dengan aspek lainnya sehingga terbentuk 12 variabel yang mempengaruhi kepuasan pelanggan tersebut, sedangkan untuk semua variable kepuasan konsumen secara umum yang di antaranya adalah *Tangible, Assurance, Responsive, Reliability* dan *Empathy* akan menghasilkan akurasi sebesar 88.00% kepuasan terhadap pelanggan tersebut.

Pada penelitian ini menguji keakuratan klasifikasi terhadap kepuasan pelanggan iklan pada Lampung Post yaitu dengan menggunakan algoritma Decision Tree, setelah itu Decision Tree dengan Forward Selection sebagai fitur seleksi. Penelitian ini menggunakan dataset privat yang diambil dari pengolahan data kuisioner yang di sebar ke pelanggan iklan yang berkunjung ke Lampung Post. Dataset yaitu dataset kepuasan pelanggan yang memiliki 200 *record* dan 12 variabel.

### 3.1. Karakteristik responden

#### a. Jenis Kelamin

Diketahui bahwa responden terbanyak adalah pria yaitu sebanyak 143 orang (71.5%), sedangkan responden wanita sebanyak 57 orang (28.5%). Hal ini dikarenakan pria lebih aktif dibandingkan wanita, sehingga berpotensi untuk memasang iklan.

#### b. Umur

Diketahui bahwa pada umur responden terbanyak adalah yang berumur >50 tahun yaitu sebanyak 46% atau sebanyak 92 respon. Hal ini disebabkan karena pada umur tersebut biasanya memiliki keinginan untuk mengetahui tingkat pencitraan dan keperluannya tersendiri.

#### c. Pekerjaan

Pekerjaan terbanyak responden adalah PNS sebanyak 22% atau 44 respon. Hal ini memberikan penjelasan bahwa responden adalah kalangan orang yang sudah bekerja di sektor Pemerintahan atau PNS merupakan konsumen yang potensial pada produk iklan Lampung Post.

#### d. Jenis Iklan

Jenis iklan terbanyak adalah iklan baris sebanyak 71% atau 142 respon. Hal ini memberikan penjelasan bahwa responden terhadap iklan baris lebih di minati di bandingkan dengan iklan kolom.

#### e. Pendapatan

Pendapatan sebagian responden adalah > 2.000.000 sebanyak 39% atau sebanyak 78 respon. Hasil tanggapan responden pada variabel *tangible* menunjukkan sebanyak 107 respon memberikan tanggapan puas terhadap variabel ini, dengan nilai indeks yang tinggi yaitu sebesar 84,5%. Alasan yang diungkapkan responden mengenai kepuasan terhadap *tangible* adalah Fasilitas pelengkap yang di sediakan Lampung post (kursi, meja, tempat parkir, dll) bersih dan dalam keadaan baik hal ini dilihat nilai indeks yang tinggi yaitu sebesar 169.

Hasil tanggapan variabel *reliability* menunjukkan sebanyak 131 respon memberikan tanggapan puas terhadap variabel ini, dengan nilai indeks yang tinggi yaitu sebesar 84,4%. Alasan yang diungkapkan responden mengenai kepuasan terhadap *reliability* adalah Lampung post memberikan perhatian serius terhadap anda ketika kesusahan dalam memilih disan iklan hal ini dilihat nilai indeks yang tinggi yaitu sebesar 168,6.

Pada tanggapan variabel *responsiveness* menunjukkan sebanyak 131 respon memberikan tanggapan puas terhadap variabel ini, dengan nilai indeks yang tinggi yaitu sebesar 84,8%. Alasan yang diungkapkan responden mengenai kesetujuannya terhadap *responsiveness* adalah Ketika anda membutuhkan sesuatu karyawan Lampung Post selalu tanggap untuk membantu anda hal ini dilihat nilai indeks yang tinggi yaitu sebesar 168,6.

Hasil tanggapan variable *assurance* menunjukkan sebanyak 123 respon memberikan tanggapan puas terhadap variabel ini, dengan nilai indeks yang tinggi yaitu sebesar 86%. Alasan yang diungkapkan responden terhadap jaminan yang diberikan adalah Lampung post selalu menjaga kebersihan dan tampilan iklan yang selalu inovatif, hal ini dapat dilihat dari nilai indeks yang tinggi yaitu sebesar 172.

Sedangkan hasil tanggapan variabel *empathy* menunjukkan sebanyak 125 respon memberikan tanggapan puas terhadap variabel ini, dengan nilai indeks yang tinggi yaitu sebesar 86,1%. Alasan yang diungkapkan adalah terhadap index yang tinggi sebesar 172,2 responden setuju terhadap variabel *empathy* adalah Karyawan Lampung post member perhatian individual pada anda.

### 3.2. Proses perhitungan analisis kepuasan pelanggan

Proses perhitungan analisis kepuasan pelanggan terhadap pelayanan iklan pada Lampung Post menggunakan algoritma C4.5 adalah sebagai berikut:

- Menghitung jumlah kasus, jumlah kasus untuk respon puas, jumlah kasus untuk keputusan respon tidak puas, dan *entropy* dari semua kasus dan kasus dibagi berdasarkan atribut *reliability*, *responsiveness*, *assurance*, *empathy*, dan *tangibles*. Setelah itu, lakukan perhitungan *gain* untuk setiap atribut.
- Diketahui jumlah kasus (S) adalah 200, jumlah respon tidak puas (S<sub>1</sub>) adalah 89, dan jumlah respon puas (S<sub>2</sub>) adalah 111. Perhitungan *entropy* dapat dihitung dengan menggunakan persamaan sebagai berikut:

$$Entropy (Total) = \left(-\frac{89}{200} * \log_2 \left(\frac{89}{200}\right)\right) + \left(-\frac{111}{200} * \log_2 \left(\frac{111}{200}\right)\right)$$

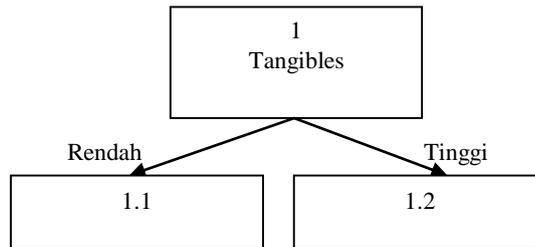
$$Entropy (Total) = 0.9912$$

- Sementara itu, nilai *gain* pada baris *reliability* dihitung dengan menggunakan persamaan sebagai berikut:

$$Gain(Total, Reliability) = Entropy(Total) - \sum_{i=1}^n \frac{|Reliability_i|}{|Total|} * Entropy(Reliability)$$

$$Gain(Reliability) = 0.9912 - \left( \left( \frac{93}{200} * 0.5836 \right) + \left( \frac{107}{200} * 0.4165 \right) \right) Gain(Reliability) = 0.4969$$

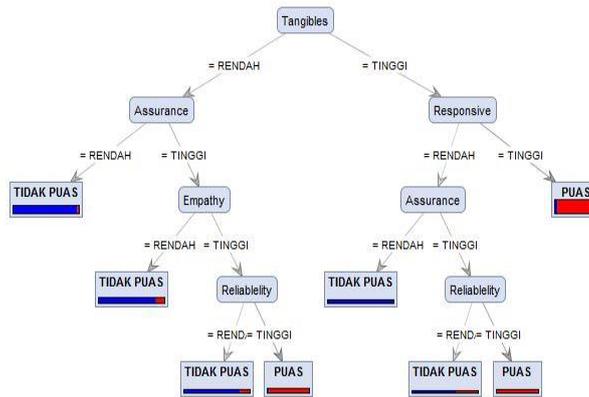
- d. Untuk perhitungan pada atribut-atribut berikutnya sama seperti perhitungan pada atribut *Tangibles*. diketahui bahwa atribut dengan *gain* tertinggi adalah *Tangibles*, yaitu sebesar 0.4969. dengan demikian *Tangibles* dapat menjadi *node* akar. Ada dua nilai atribut dari *Tangibles* yaitu rendah dan tinggi. Dari kedua nilai atribut tersebut masih memerlukan perhitungan lebih lanjut. Dan hasil perhitungan tersebut dapat digambarkan pohon keputusan seperti pada Gambar 2.



Gambar 2. Pohon keputusan *node* 1

**3.3. Pengujian**

- a. Pengujian dengan menggunakan *Decision Tree C4.5*



Gambar 3. Pohon Keputusan hasil analisis

Dari gambar di atas terlihat dari dua belas variable yang terbentuk dari dataset yang memiliki 200 record dapat terlihat *root note* utama yang mewakili keduabelas aspek tersebut yaitu *Tangibles* sebagai aspek yang paling berpengaruh di antara seluruh aspek atau variable yang terbentuk dari dataset tersebut.

Akan tetapi dari keduabelas aspek atau variable yang terbentuk dari dataset tersebut, dimana tingkat keakurasiannya menurun hanya 84,50% dari keseluruhan aspek yang ada.

|  |                 |           |                 |
|--|-----------------|-----------|-----------------|
| accuracy: 84.50% +/- 9.34% (mikro: 84.50%) |                 |           |                 |
|  | true TIDAK PUAS | true PUAS | class precision |
| pred. TIDAK PUAS                           | 73              | 15        | 82.95%          |
| pred. PUAS                                 | 16              | 96        | 85.71%          |
| class recall                               | 82.02%          | 86.49%    |                 |

Gambar 4. Akurasi dari C4.5

Dari Gambar 4 menunjukkan tingkat keakurasian data dalam tingkat kepuasan konsumen terhadap pelayanan iklan Lampung Post apabila seluruh aspek atau variabel diikuti sertakan dalam model *decision tree* C4.5.

b. Pengujian Menggunakan *Decision Tree* dan di optimasikan dengan *Feature Selection Forward*

Berikut adalah gambar pengujian dengan menggunakan *Feature Selection Forward* untuk meningkatkan akurasi dataset yang terdiri dari duabelas aspek atau variabel dan memiliki 200 *record* atau sebanyak 200 respon dari konsumen Lampung Post.

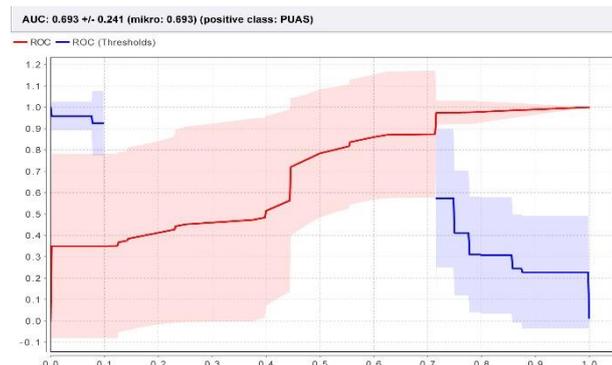
| accuracy: 90.00% +/- 5.92% (mikro: 90.00%) |                 |           |                 |
|--|-----------------|-----------|-----------------|
|  | true TIDAK PUAS | true PUAS | class precision |
| pred. TIDAK PUAS                           | 76              | 7         | 91.57%          |
| pred. PUAS                                 | 13              | 104       | 88.89%          |
| class recall                               | 85.39%          | 93.69%    |                 |

Gambar 5. Akurasi *Feature Selection Forward*

Dari Gambar 5 tersebut terlihat dari 12 variabel dan 200 record yang terbentuk dari dataset, terlihat bahwa tingkat keakurasian dapat di tingkatkan dengan menggunakan model *Feature Selection Forward* sebesar 90%.

c. Evaluasi dengan kurva ROC

Berikut adalah kurva AUC dengan algoritma *Decision Tree* dengan menggunakan dataset privat yang telah terbentuk yang di optimasikan dengan *Feature Selection Forward*.



Gambar 6. AUC 12 Aspek

Pada Gambar 6. menunjukkan grafik ROC dengan nilai AUC (*Area Under Curve*) sebesar 0.69 dengan tingkat diagnosa *Poor Classification* untuk model *Decision Tree Gain Ratio*. Dengan menggunakan 12 aspek atau variabel dari dataset yang ada.

d. Tahap eksperimen

Dengan menerapkan algoritma *Decision Tree* dengan menggunakan model yakni *gain ratio* dengan nilai *cross validation* mulai 10, 20 dan 30. Hasil yang di dapat dari eksperimen tersebut sebagai berikut.

Tabel 1. Kreteria 10,20,30 *validation* dan *Gain Rasio*

| <i>Validation</i> | <i>Gain Rasio</i> |
|-------------------|-------------------|
| 10                | 84,50%            |
| 20                | 87,50%            |
| 30                | 87,94%            |

Dari Tabel 1 penerapan algoritma Decision Tree tanpa seleksi fitur dari model kriteria gain ratio, menunjukkan bahwa semakin tinggi Validation maka akan mempengaruhi peningkatan data dengan 30 *fold cross validation* dapat mencapai peningkatan sampai 87,94%.

e. Tahapan *feature delection forward*

Pada tahapan ini dataset dapat di tingkatkan menjadi lebih baik lagi dengan *cros validation* antara 10, 20 dan 30 terdapat perubahan yaitu peningkatan akurasi yang lebih baik lagi, peningkatan tersebut dapat mencapai 90%.

#### 4. KESIMPULAN DAN SARAN

##### 4.1. Kesimpulan

Berdasarkan hasil eksperimen, mulai tahap awal hingga evaluasi, dapat ditarik kesimpulan bahwa:

1. Algoritma *Decision Tree* yang di terapkan untuk menganalisis kepuasan pelanggan iklan pada Lampung Post terletak pada aspek *Tangible*, dimana aspek atau variabel tersebut menjadi *root note* utama dalam diagram pohon yang terbentuk pada saat pengujian dengan menggunakan RapidMiner, dengan menggunakan dataset berdasarkan dari 12 aspek/variabel yang terdiri dari 200 *record* yang terbentuk variabel *Tangible* yang dominan menjadi akar pohon dalam penerapan metode *Decision Tree* tersebut. Adapun aspek *Tangible* tersebut terdiri dari:
  - a. Ruang *lobby* yang bersih, nyaman dan rapi
  - b. Lokasi Lampung post mudah di temukan
  - c. Fasilitas pelengkap yang di sediakan Lampung post (kursi, meja, tempat parkir, dll) bersih dan dalam keadaan baik
  - d. Karyawan Lampung Post berpenampilan rapih dan bersih
2. Algoritma *Decision Tree* C4.5 dan *Feature Selection Forward* dengan menggunakan model kriteria *gain ratio*, hasil akurasi yang paling tinggi terdapat pada model *gain ratio* yakni 84.50% terhadap 12 aspek yang ada pada saat sebelum di reduksi dengan menggunakan *Feature Selection Forward*. Sedangkan padasaat di reduksi dengan menggunakan *Feature Selection Forward* pada 12 aspek atau variabel yang terbentuk dari dataset mengalami peningkatan akurasi sebesar 90%.
3. Selanjutnya model kriteria gain ratio dari 12 aspek atau variabel yang terbentuk hasil akurasinya menjadi 90.00. Serta memiliki nilai rata-rata AUC 0,80-0,90 dan ini termasuk klasifikasi baik.
  - a. Dengan adanya analisis kepuasan pelanggan terhadap pelayanan iklan Lampung Post menggunakan algoritma Decision Tree dan *Feature Selection Forward*, maka tingkat kepuasan konsumen iklan Lampung Post dapat terukur dengan jelas.
  - b. Setelah dilakukan analisis, dari beberapa aspek yang menjadi tolak ukur penilaian kepuasan pelanggan maka dapat diketahui bahwa aspek *Tangible* yang paling dominan dari beberapa aspek kualitas pelayanan yang diberikan oleh pihak Lampung Post.

##### 4.2. Saran

Setelah melakukan analisis kepuasan pelanggan terhadap pelayanan iklan Lampung Post menggunakan algoritma *Decision Tree* dan *Feature Selection Forward* maka ada beberapa saran yang diberikan penulis, antara lain:

1. Metode *Decision Tree* dan *Feature Selection Forward* merupakan metode pada data mining yang dapat digunakan untuk memprediksi dan mengklasifikasi di berbagai bidang sehingga masih terbuka bagi peneliti lain untuk menggunakan metode tersebut pada permasalahan prediksi dan klasifikasi penerapan data mining di bidang lainnya.
2. Dalam penelitian ini menggunakan metode *Decision Tree* dan *Feature Selection Forward* sebagai peningkatan akurasinya, maka masih memungkinkan dilakukan penelitian lebih lanjut terhadap analisis kepuasan pelanggan terhadap pelayanan iklan Lampung Post, dengan menggunakan metode lainnya untuk lebih meningkatkan keakurasiannya.

#### ACKNOWLEDGEMENTS

Paper ini adalah penelitian dari Tesis saya dan untuk pembiayaannya sendiri dari biaya sendiri.

**DAFTAR PUSTAKA**

- [1] Arya Purnanditya, Bondhan. 2014. “Penerapan Fitur Seleksi Forward Selection Menggunakan Algoritma Naive Bayes Untuk Menentukan Atribut Yang Berpengaruh Pada Klasifikasi Kelulusan Mahasiswa Universitas AKI Semarang”. Jurnal Ilkom.
- [2] Cahya Putri Buani, Duwi. 2016. “Optimasi Algoritma Naive Bayes dengan Menggunakan Algoritma Genetika untuk Prediksi Kesuburan (Fertility) Program Studi Teknik Komputer”. Jurnal Evolusi Volume 4 Nomor 1.
- [3] Defiyanti, Sofi. 2013. “Analisis dan Prediksi Kinerja Mahasiswa Menggunakan Teknik Data Mining”. Jurnal Ilkom Syntax Vol. 2 Ed. 1. Tahun 2013.
- [4] Dennis Aprilla C, Donny Aji Baskoro, Lia Ambarwati, I Wayan Simri Wicaksana, 2013. “Belajar Data Mining dengan RapidMiner”. E-paper: Jakarta.
- [5] Fatchur Rohman, Ibnu. 2015. “Penerapan Algoritma C4.5 Pada Kepuasan Pelanggan Perum DAMRI”. Jurnal Ilkom.