

# Implementasi Data Mining Dalam Pemilihan Laptop Berbasis Algoritma C4.5 Pada Software WEKA

<sup>1</sup>Baginda Harahap

[profesionalbaginda@gmail.com](mailto:profesionalbaginda@gmail.com)

## ABSTRAK

Kemajuan globalisasi saat ini memberikan dampak yang kuat di berbagai bidang teknologi. Laptop atau komputer *portabel* adalah komputer kecil dan ringan sehingga dapat membawanya ke mana saja tentu tergantung pada ukuran, bahan dan spesifikasi laptop tentu menjadi prioritas utama yang sangat dipertimbangkan sebelum membelinya. Penggunaan laptop sama seperti komputer yang dapat membantu kita untuk menyelesaikan tugas, melakukan pekerjaan dengan baik dan cepat. Orang terkadang punya laptop banyak permainan untuk memberikan hiburan. Toko Jaya Utama Medan adalah toko yang menawarkan berbagai macam laptop. Terlalu banyak jenis laptop membuat konsumen sulit untuk memilih laptop mana yang akan dipilih karena sulit memberikan informasi dari pihak toko jelas bagi konsumen. Kemudian juga membutuhkan teknik data *mining* C4.5 Algoritma dimana teknik ini mengklasifikasikan pemilihan laptop yang baik berdasarkan variabel harga, jumlah RAM, prosesor, ukuran, daya tahan baterai. Kebenaran data yang diproses diperiksa di aplikasi WEKA dengan menggunakan Algoritma C4.5. Hasil ini diperoleh dengan mencari nilai *entropi*. Dan keuntungan dari semua variabel menghasilkan pohon keputusan. Maka dengan hal itu ada tiga variabel yang sangat mempengaruhi pemilihan laptop, yaitu harga, prosesor dan masa pakai baterai. Untuk total 80 data yang dipilihkan untuk memilih laptop yang benar hanya 27, yang salah 53.

**Kata Kunci:** Klasifikasi, Algoritma C4.5, Data *Mining*, WEKA

## ABSTRACT

The current progress of globalization has a strong impact on various fields of technology. A laptop or portable computer is a small and light computer, so you can take it anywhere, of course, depending on the size, material and specifications of the laptop, of course, this is a top priority to consider before buying it. The use of laptops is the same as computers that can help us complete tasks, do work well and quickly. People sometimes have laptops with lots of games to provide entertainment. Toko Jaya Utama Medan is a store that offers a wide range of laptops. Too many types of laptops make it difficult for consumers to choose which laptop to choose because it is difficult to provide clear information from the store for consumers. Then it also requires a data mining technique C4.5 Algorithm where this technique classifies the selection of a good laptop based on price variables, amount of RAM, processor, size, battery life. The correctness of the processed data is checked in the WEKA application using the C4.5 Algorithm. This result is obtained by finding the entropy value. And the advantages of all the variables generate a decision tree. So with that in mind there are three variables that greatly influence laptop selection, namely price, processor and battery life. For a total of 80 selected data to choose the correct laptop only 27, 53 wrong.

**Keywords:** Classification, C4.5 Algorithm, Data Mining, WEKA

## PENDAHULUAN

Perkembangan globalisasi saat ini sangat mempengaruhi berbagai bidang, salah satunya yaitu bidang teknologi. Perkembangan teknologi yang semakin pesat dan memasuki *fase society* 5.0. Dari sisi kehidupan, teknologi juga sangat dibutuhkan untuk dapat membantu dan mempermudah pekerjaan. Barang-barang elektronik pun tidak lagi menjadi barang mewah saat ini. Dampak positif perkembangan teknologi di *fase society* 5.0 yakni menjadikan manusia semakin cerdas, individual, serta memiliki harapan yang tinggi karena dengan mudah terhubung langsung ke dunia global yang menyajikan data-data terbaru. Harapan tersebut akan terlihat ketika manusia dilihat dari sisi konsumtifnya. Mereka akan mencari barang yang paling cocok dengan kriteria mereka, untuk memenuhi kebutuhan seiring perkembangan zaman. Barang-barang elektronik juga telah mengalami perkembangan seperti televisi, telepon, komputer dan laptop. Laptop atau komputer tentang merupakan sebuah komputer yang dapat dibawa kemana saja serta memiliki ukuran kecil dan ringan, tergantung pada ukuran, bahan, dan spesifikasi laptop tersebut. Laptop juga sering disebut dengan komputer *notebook* atau *notebook* saja. Laptop mempunyai kegunaan mirip dengan komputer desktop (*Desktop Computers*) seperti umumnya yaitu dapat memudahkan mengerjakan tugas dan pekerjaan dengan kilat dan baik. (Ginting, 2015).

Disamping itu juga dengan adanya laptop dapat digunakan untuk menambah penghasilan dengan banyaknya penjualan secara *online*. Bagi kalangan anak muda, ada kalanya untuk hiburan dengan banyaknya game yang ditawarkan. Oleh sebab itu, Laptop bukan lagi menjadi barang mewah. Laptop juga sudah menjadi aspek penting dari kebutuhan berbagai kalangan seperti karyawan, pembisnis, dan siswa maupun mahasiswa untuk menunjang pekerjaan maupun studinya. Toko Jaya Utama yang beralamat di Jl. Merak Jingga Kota Medan ini adalah salah satu toko yang menyediakan berbagai merk laptop. Toko Jaya Utama termasuk toko yang paling ramai dikunjungi oleh banyak orang yang akan membeli maupun sekedar tanya-tanya laptop. Sehingga Toko Jaya Utama sering kali mengalami kesulitan dalam membantu konsumen menentukan jenis dan tipe laptop yang akan dibeli. Hal ini terjadi karena banyaknya jenis dan harga bersaing yang ditawarkan oleh pihak produsen laptop. Untuk memilih laptop yang tepat sesuai dengan kebutuhan dan dana yang disiapkan oleh konsumen, maka memerlukan suatu analisa yang akurat yang dapat dijadikan bahan pertimbangan dari banyaknya kriteria dan faktor. Kriteria - kriteria yang dapat dijadikan acuan dalam memilih suatu laptop antara lain yaitu harga, kapasitas RAM, prosesor, ukuran dan daya tahan baterai dari laptop tersebut. Untuk mengolah semua kriteria tersebut maka diperlukan suatu sistem yang dapat menentukan dan mengelompokkan kriteria- kriteria yang dipilih oleh konsumen, sehingga dapat membantu konsumen dalam memilih laptop yang tepat. *Data mining* merupakan suatu proses pengolahan data dengan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar. *Data mining* juga diartikan sebagai proses untuk membangun hubungan, struktur dan model baru yang dapat digunakan dalam penyeleksian data yang sangat besar. Salah satu teknik dari *data mining* adalah klasifikasi.

(Seng Hansun, 2014) *Algoritma* C4.5 merupakan *algoritma* yang banyak dimanfaatkan dalam proses pengklasifikasian pohon keputusan sebab mempunyai kelebihan utama dibandingkan *algoritma* lainnya. Kelebihan *algoritma* C4.5 dapat menghasilkan pohon keputusan yang mudah diinterpretasikan, mempunyai angka kebenaran yang bisa diterima, efisien dalam menangani atribut bertipe diskret dan dapat menangani atribut bertipe diskret dan numerik. (Seng Hansun, 2014) Pengolahan data mining menggunakan metode *algoritma* C4.5 yang bertujuan untuk memperoleh informasi berupa pemilihan laptop di Toko Jaya Utama Medan. *Data mining* memanfaatkan prosedur khusus untuk memperoleh skema maupun berita merampas atas sebuah berita. Tahapan dari *Data mining* yang digunakan dalam mengekstrak serta mengidentifikasi suatu informasi yang dibutuhkan oleh *database* besar dengan cara teknik statistik, matematika, kecerdasan buatan, dan *machine*. *Data mining* merupakan sebuah cara yang dibantu oleh metode statistik matematika, kecerdasan buatan, dan *machine learning* untuk menghasilkan dan membuktikan berita yang bermanfaat serta ilmu yang berhubungan dengan *database* besar.

## TINJAUAN PUSTAKA

### Knowledge Discovery In Database (KDD)

*Knowledge Discovery in Database* (KDD) adalah merupakan metode dalam mendapatkan keahlian basis data yang ada. Dalam *database* diperoleh grafik-grafik yang saling terhubung satu sama lainnya. Pengambilan keputusan dari hasil keahlian yang didapatkan melalui tahapan yang berguna sebagai basis pengetahuan (*knowledge base*). Istilah *Knowledge Discovery in Database* (KDD) dan *data mining* biasanya dimanfaatkan secara bergantian dalam memperjelas tahapan penggalian informasi tertanam pada *database* besar. Dari kedua perumpamaan tersebut mempunyai ide tidak serupa namun memiliki keterikatan diantaranya, dan satu diantaranya yaitu bagian dari keseluruhan prosedur *Knowledge Discovery in Database* (KDD) adalah *data mining*. (Mardi, 2018) Tahapan *Knowledge Discovery in Database* (KDD) secara ilustrasi bisa dijelaskan sebagai berikut:

#### 1. Data Selection

Sebelum proses penggalian informasi *Knowledge Discovery in Database* (KDD) dimulai, maka diperlukan proses penyeleksian data dari sekelompok data operasional. Hasil dari penyeleksian data tersebut disimpan kedalam berkas terpisah dari database operasional untuk proses data *mining*.

#### 2. Pre-processing / Cleaning

Tahapan pembersihan data pada saat fokus *Knowledge Discovery in Database* (KDD) dilaksanakan sebelum proses data *mining* dimulai. tahapan penyortiran data meliputi yaitu penyortiran data ganda, mendalami inkonsisten data, dan mengubah data yang salah cetak. Selain itu juga ada proses *enrichment*, adalah data atau informasi luar yang dibutuhkan oleh *Knowledge Discovery in Database* (KDD) yang valid melalui tahapan memperbanyak data yang telah diketahui sebelumnya.

#### 3. Transformation Coding

Tahapan *coding* dalam *Knowledge Discovery in Database* (KDD) ialah prosedur yang bergantung pada macam atau bentuk dari informasi yang akan dicari pada *database* serta prosedur transmudasi atas berita yang sudah diambil berdasarkan kemauan pada data *mining*.

#### 4. Data Mining

Data *mining* merupakan suatu cara dalam mencari sketsa serta berita menarik dalam yang telah terpilih oleh teknik dan metode tertentu. Data *mining* memiliki banyak teknik dan metode atau *algoritma* yang beranekragam sehingga pemilihan metode dan *algoritma* yang tepat sangat bergantung pada hasil dan tahapan dari *Knowledge Discovery in Database* (KDD) secara keseluruhan.

#### 5. Interpretation / Evaluation

Bentuk informasi dari proses data mining yang telah diperoleh yaitu harus memiliki bentuk yang sederhana yang gampang dipahami oleh pihak yang bersangkutan. Proses ini adalah bagian dari *Knowledge Discovery in Database* (KDD) yang diartikan sebagai *interpretation*. proses ini mencakup pemeriksaan apakah sketsa atau berita yang didapatkan bertolak belakang dengan fakta atau hipotesis sebelumnya yang sudah ada.

### Data Mining

Data mining adalah suatu proses pengerukan atau pengumpulan informasi penting dari suatu data yang besar. Proses data mining seringkali menggunakan metode statistika, matematika, hingga memanfaatkan teknologi artificial intelligence. Nama alternatifnya yaitu Knowledge discovery (mining) in databases (KDD), knowledge extraction, data/pattern analysis, data archeology, data dredging, information harvesting, business intelligence, dan lain-lain. Jika di lihat di lihat pada gambar dalam proses KDD tersebut, Banyak konsep dan teknik yang di gunakan dalam proses data mining. Proses tersebut membutuhkan beberapa langkah untuk mendapatkan sebuah data yang di inginkan. Dalam proses KDD tersebut termasuk melakukan pembersihan data, integrasi data, seleksi data, transformasi, penambangan data, evaluasi pola, dan presentasi pengetahuan.

### Algoritma C4.5

Salah satu algoritma dalam membentuk suatu pokok kesimpulan yang biasa dipakai adalah

*algoritma* C4.5. *Algoritma* C4.5 adalah pengembangan atas ID3, pengembangannya untuk melumpuhkan nilai yang hilang, biasanya melumpuhkan data selanjutnya serta harapan. *Algoritma* C4.5 merupakan metode untuk menstruktur sebuah pohon keputusan berkenaan dengan berita yang didapatkan terlebih dahulu. Gaya pertama yang melandasi *algoritma* ini adalah pengstrukturian potoh kesimpulan berkenaan gaya cabang yang memiliki marka *gain* tertinggi berkenaan dengan marka *entropy* cabang tersebut semisal *under step* dari *algoritma* ini yaitu mempunyai nilai *gain* tertinggi berdasarkan nilai *entropy* bagian dari proses pengklasifikasian. Rumus *gain* tertinggi dapat dilihat sebagai berikut ini (Elisa & Ervina, 2018).

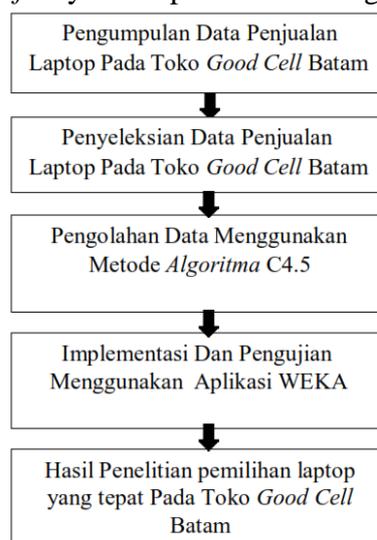
### **Waikato Environment for Knowledge Analysis (WEKA)**

*Waikato Environment for Knowledge Analysis* adalah sebuah kelompok alat pembelajaran mesin praktis yang disingkat WEKA. Aplikasi ini diciptakan di Universitas Waikato, New Zealand bagi ulasan, keahlian dan macam-macam aplikasi. Keuntungan menggunakan WEKA yaitu khususnya klasifikasi yang dilandasi oleh pendekatan-pendekatan sangat mampu menyelesaikan masalah-masalah data *mining* dalam dunia nyata. Selain itu WEKA juga sangat gampang diterapkan diberberapa tahapan berbeda dan WEKA menyediakan alat untuk *pre processing* data, klasifikasi, regresi, *clustering*, aturan asosiasi, dan visualisasi. Ada beberapa metode dari data mining yang bisa dieksekusi menggunakan aplikasi WEKA yang sering digunngakan yaitu pohon keputusan (*decision tree*). Kontruksi dai pohon keputusan yang melibatkan *node* keputusan yang akan disambungkan ke setiap bagian-bagian mulai simpul akar sampai ke *node* daun (akhir) merupakan bagian yang menarik dari metode klasifikasi. Pada *node* kesimpulan indikasi bakal dites, lalu tiap-tiap reaksi atas memperoleh bagian. Tiap-tiap bagian akan ditunjukkan ke *node* lain atau ke *node* akhir untuk mendapatkan hasil yaitu berupa keputusan. (Pujiono, Amborowati, & Suyanto, 2013).

## **METODE PENELITIAN**

### **Desain Penelitian**

Sebelum masuk kebagian teknik penelitian, desain penelitian perlu dirancang terlebih dahulu. Dengan tujuan untuk memudahkan penelitian ke tahapan selanjutnya. Adapun Desaian penelitian dalam penelitian ini mulai dari awal hingga akhir, dan desain penelitian ini bisa kita sebut metode *waterfall* yaitu dapat dilihat sebagai berikut:



**Gambar 1. Desain dan Metode Penelitian**

## HASIL DAN PEMBAHASAN

### 1. Pengumpulan Data

Proses ini dilakukan guna mendapatkan berita yang dibutuhkan dalam penelitian guna mencapai tujuan yang diinginkan oleh peneliti. Data yang diperoleh yaitu berupa data penjualan laptop Toko Jaya Utama, melalui wawancara, observasi dan *study* pustaka mengenai data *mining* serta metode *Algoritma C4.5*.

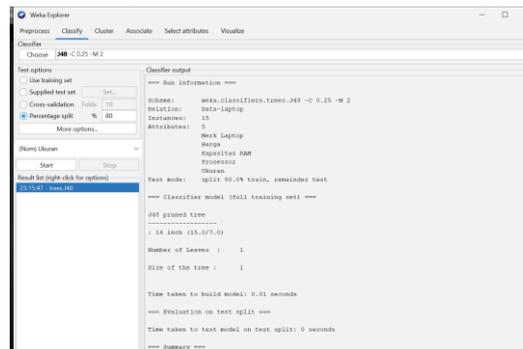
**Tabel 1. Penjualan Laptop di Toko Jaya Utama**

No	Merk Laptop	Harga	Kapasitas RAM	Processor	Ukuran
1	Acer Aspire 3 Slim A315	4.899.000,-	4GBDDR4	Intel Celeron N5100	15,6 inch
2	Lenovo IdeaPad Slim D330 Flex	5.040.000,-	8GB	Intel Celeron N4020	10.1 inch
3	ASUS Vivo Book 15 A516MAO N4020	4.599.000,-	4GB DDR4	Intel® Celeron® N4020 Processor 1.1	15,6 inch
4	Lenovo IdeaPad Slim 3i 14IGL05	4.229.000,-	4GB DDR4	Intel Celeron N4020	14 inch
5	HP Laptop 14s dq0508TU	5.049.000,-	4GB DDR4	Intel Celeron N4120	14 inch
6	Xiaomi RedmiBook 15	6.999.000,-	DDR4 8 GB 3200MHz	11th Generation Intel® Core™ i3-1115G4	15,6 inch
7	Lenovo V14 G2 ITL	3.699.000,-	4GB Soldered DDR4	Intel Core i3-1115G4	14 inch
8	ASUS Vivo Book 14 A416FA	5.799.000,-	4GB DDR4 on board	Intel® Core™ i3-10110U Processor 2.1 GHz	14 inch
9	ASUS Vivo Book Go 14 Flip TP1400KA	8.878.000,-	8GB DDR4 on board	Intel® Pentium® Silver N6000	14 inch
10	Acer Aspire 5 Slim A514	6.549.000,-	4GB DDR4 onboard	Intel Core i3-1115G4	14 inch
11	HP Laptop 14s dq2614TU	6.299.000,-	4 GB DDR4-2666 SDRAM	11th Generation Intel® Core™ i3 processor	14 inch

No	Merk Laptop	Harga	Kapasitas RAM	Processor	Ukuran
12	Dell Vostro 3401	6.899.000,-	4GB	Intel Core™ i3-1005G1	14 inch
13	Dell XPS 13	10.500.00,-	LPDDR3 8GB 2133MHz	Intel® Core™ i7-8250U Processor	13,3 inch
14	Acer Chromebook Spin 713	5.699.000,-	8GB	Intel Core i5-1135G7	13,5 inch
15	HP Pavilion Aero 13	11.999.000,-	8 GB DDR4-2666 SDRAM	AMD Ryzen™ 5 processor	13,5 inch

## 2. Penyeleksian Data

Proses ini dilakukan guna menghasilkan data yang benar-benar sesuai dengan kebutuhan serta tujuan penelitian yang nantinya akan dijadikan sampel. Data yang dipilih yaitu berupa data-data Penjualan Laptop Pada Toko Jaya Utama dari tahun 2021 hingga 2022 Per Agustus.

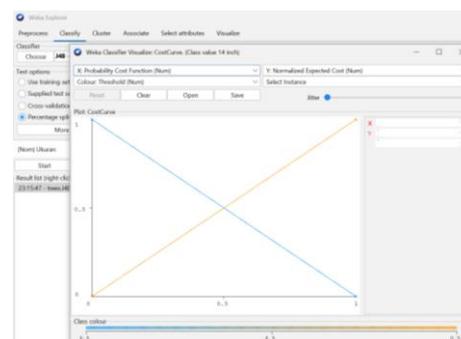


Gambar 2. Clasification WEKA Terhadap Data Penjualan Laptop

Berdasarkan klasifikasi weka terhadap data penjualan laptop Toko Jaya Utama, maka ada 5 atribut yang menjadi klasifikasi yaitu merk laptop, harga, kapasitas ram, prosessor dan ukuran layer. Maka dengan klasifikasi weka dengan *test mode* 80% adalah 14 inch yaitu menandakan kebutuhan ukuran layer.

## 3. Visualize CostCurve

Aplikasi WEKA akan digunakan sebagai aplikasi yang digunakan dalam pengujian untuk pemilihan laptop pada Toko Jaya Utama dengan menggambarkan visualize CostCurve sebagai berikut:

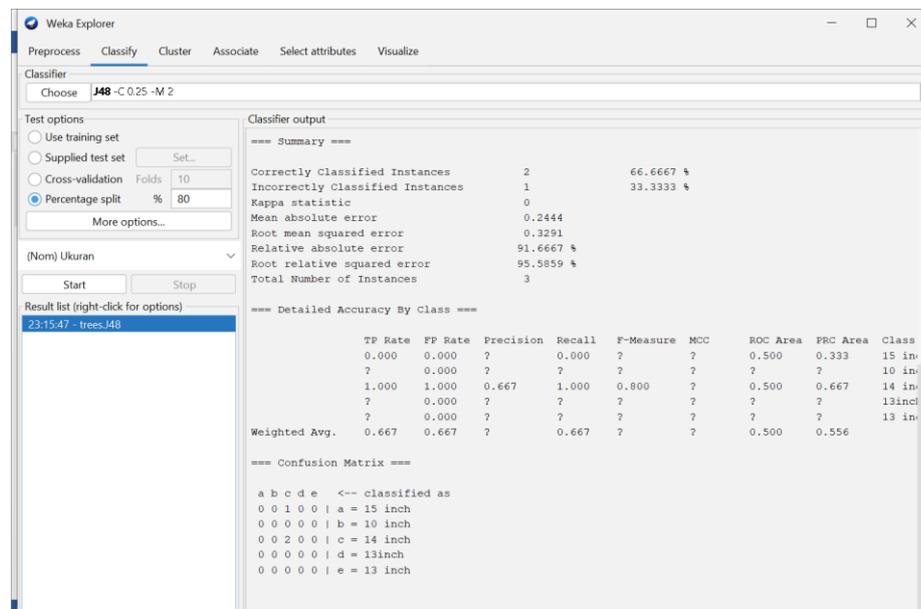


Gambar 3. Visualize CostCurve 14 inch

Berdasarkan gambar visualize yang ditampilkan di WEKA, bahwa yang menjadi pilihan konsumen untuk membeli laptop itu adalah ukuran layer, yaitu 14 inch.

#### 4. Hasil Penelitian

Hasil yang menjadi tujuan dari penelitian ini adalah untuk menentukan pemilihan laptop pada Toko Jaya Utama yang telah diolah dengan teknik data *mining* menggunakan metode *algoritma* C4.5 dengan proses pengujian menggunakan aplikasi WEKA.



Gambar 4. Classifier Output Weka Penjualan Laptop

Berdasarkan hasil classifier output weka maka ada 2 cluster yang paling banyak diminati pada pembelian laptop di Toko Jaya Utama yaitu terdapat 2 orang memilih dari sisi ukuran layer 14 inch, 15 inch = 1, layer 10 inch = 0, layer 13 inch = 0.

#### 5. Teknik Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini menggunakan 3 teknik yaitu, observasi, wawancara, dan studi pustaka. Penjelasan dari ketiga teknik pengumpulan data yaitu sebagai berikut:

##### a. Observasi

Observasi merupakan serangkaian proses psikologis dan biologis yang tersusun secara kompleks. Observasi dalam penelitian ini dilakukan secara langsung pada Toko Jaya Utama untuk melakukan pengamatan terhadap data penjualan laptop serta mengetahui variabel-variabel yang dibutuhkan dalam proses pengolahan data sebagai bahan dalam penelitian.

##### b. Wawancara

wawancara merupakan suatu proses untuk memperoleh informasi langsung dari sumbernya. Dalam penelitian ini peneliti melakukan wawancara langsung dengan karyawan Toko Jaya Utama yang bertujuan untuk memperoleh informasi tentang penjualan laptop serta kendala-kendala yang sering mereka alami saat memberikan informasi yang tepat bagi konsumen saat membeli laptop ditoko Toko Jaya Utama.

##### c. Studi Pustaka

Pada penelitian ini, peneliti melakukan pengelompokkan beberapa sumber sebagai partisan penelitian yang diperoleh dari buku-buku, jurnal dan dari bermacam referensi yang memiliki keterikatan dengan objek penelitian.

## 6. Operasional Variabel

Proses pengolahan data dilakukan sesuai dengan data penjualan laptop pada Toko Jaya Utama. Variabel–variabel yang diperlukan dalam proses analisa data yaitu berdasarkan:

### a. Harga

Harga merupakan aspek utama dalam membeli laptop, karena harga bisa mengukur seberapa bagus laptop yang akan dibeli.

### b. Kapasitas RAM

Kapasitas RAM merupakan salah satu faktor yang mempengaruhi kinerja laptop. Kapasitas setiap laptop tentu berbeda. Semakin besar RAM yang dipasang maka semakin besar juga tempat penyimpanan data. Oleh karena itu kapasitas RAM perlu diperhatikan saat membeli laptop.

### c. Prosesor

Bagian yang tidak kalah penting dalam pemilihan laptop ialah prosesor. Prosesor merupakan bagian dari dapur pacu. Ada beragam jenis prosesor yang dijual di pasaran baik tipe AMD maupun Intel.

### d. Ukuran Laptop

Ukuran layar yang dipilih saat membeli laptop baru, tentunya akan memengaruhi ukuran laptop.

#### 1. Daya Tahan Batrai

2. Rata-rata laptop yang dijual memiliki daya tahan baterai kisaran 6-8 jam.

3. Namun ketahanan baterai agak sulit diukur karena tergantung pemakaian *laptop*.

4. Dari variabel-variabel yang telah disebutkan dan dijelaskan diatas, maka

5. Keputusan yang akan dihasilkan yaitu **Tepat** dan **Tidak Tepat**.

## KESIMPULAN

Berkenaan dengan struktur dari perancangan sistem diatas, terlihat bahwa memperoleh data penjualan laptop dengan kriteria seperti harga, kapasitas RAM, prosesor, ukuran dan daya tahan batrai dari laptop pada Toko Jaya Utama, kemudian dilakukan proses penyeleksian data hingga memperoleh data yang layak yang akan diproses ke tahapan selanjutnya. Setelah memperoleh hasil penyeleksian data telah selesai, maka data penjualan tersebut selanjutnya diolah kedalam teknik data mining menggunakan metode algoritma C4.5, selanjutnya diuji dengan menggunakan aplikasi WEKA, untuk mengetahui keakuratan antara proses perhitungan manual tersebut sesuai dengan hasil pengujian melalui aplikasi WEKA dalam menentukan pilihan laptop yang tepat berdasarkan data yang diolah. Dan hasil dari classifier output weka yang menjadi prioritas pembeli di Toko Jaya Utama adalah beradaskan ukurannya yaitu 14 inch.

## REFERENSI

- Aher, S. B., & Lobo, L. (2011). Data mining in educational system using weka. *International Conference on Emerging Technology Trends (ICETT)*, 3, 20–25.
- Baker, R., & others. (2010). Data mining for education. *International Encyclopedia of Education*, 7(3), 112–118.
- bin Othman, M. F., & Yau, T. M. S. (2007). Comparison of different classification techniques using WEKA for breast cancer. *3rd Kuala Lumpur International Conference on Biomedical Engineering 2006: Biomed 2006, 11--14 December 2006 Kuala Lumpur, Malaysia*, 520–523.
- Bouckaert, R. R. (2004). *Bayesian network classifiers in weka*.
- Bouckaert, R. R., Frank, E., Hall, M. A., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2010). WEKA---Experiences with a Java Open-Source Project. *The Journal of Machine Learning Research*, 11, 2533–2541.
- Cabena, P., Hadjinian, P., Stadler, R., Verhees, J., & Zanasi, A. (1998). *Discovering data mining: from concept to implementation*. Prentice-Hall, Inc.
- Chen, M.-S., Han, J., & Yu, P. S. (1996). Data mining: an overview from a database perspective. *IEEE Transactions on Knowledge and Data Engineering*, 8(6), 866–883.
- Frank, E., Hall, M., Holmes, G., Kirkby, R., Pfahringer, B., Witten, I. H., & Trigg, L. (2010). Weka-

- a machine learning workbench for data mining. *Data Mining and Knowledge Discovery Handbook*, 1269–1277.
- Frank, E., Hall, M., Trigg, L., Holmes, G., & Witten, I. H. (2004). Data mining in bioinformatics using Weka. *Bioinformatics*, 20(15), 2479–2481.
- Garner, S. R., & others. (1995). Weka: The waikato environment for knowledge analysis. *Proceedings of the New Zealand Computer Science Research Students Conference, 1995*, 57–64.
- Gupta, G. K. (2014). *Introduction to data mining with case studies*. PHI Learning Pvt. Ltd.
- Hand, D. (2006). Data mining. *Encyclopedia of Environmetrics*.
- Hand, D. J. (2007). Principles of data mining. *Drug Safety*, 30, 621–622.
- Kotthoff, L., Thornton, C., Hoos, H. H., Hutter, F., & Leyton-Brown, K. (2019). Auto-WEKA: Automatic model selection and hyperparameter optimization in WEKA. *Automated Machine Learning: Methods, Systems, Challenges*, 81–95.
- Luka, P. D., Adedeji, A. J., Jambol, A. R., Ifende, I. V., Luka, H. G., Choji, N. D., Weka, R., Settyapalli, T. B. K., Achenbach, J. E., Cattoli, G., & others. (2022). Coinfections of African swine fever virus, porcine circovirus 2 and 3, and porcine parvovirus 1 in swine in Nigeria. *Archives of Virology*, 167(12), 2715–2722.
- Mannila, H. (1997). Methods and problems in data mining. *Database Theory—ICDT'97: 6th International Conference Delphi, Greece, January 8--10, 1997 Proceedings 6*, 41–55.
- Markov, Z., & Russell, I. (2006). An introduction to the WEKA data mining system. *ACM SIGCSE Bulletin*, 38(3), 367–368.
- Naik, A., & Samant, L. (2016). Correlation review of classification algorithm using data mining tool: WEKA, Rapidminer, Tanagra, Orange and Knime. *Procedia Computer Science*, 85, 662–668.
- Read, J., Reutemann, P., Pfahringer, B., & Holmes, G. (2016). *Meka: a multi-label/multi-target extension to weka*.
- Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12–27.
- Sharma, N., Bajpai, A., & Litoriya, M. R. (2012). Comparison the various clustering algorithms of weka tools. *Facilities*, 4(7), 78–80.
- Singhal, S., & Jena, M. (2013). A study on WEKA tool for data preprocessing, classification and clustering. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 2(6), 250–253.
- Sumathi, S., & Sivanandam, S. N. (2006). *Introduction to data mining and its applications* (Vol. 29). Springer.
- Weka, R. P., Ikeh, E. I., & others. (2009). Seroprevalence of cysticercosis and intestinal parasitism in pigs in Jos metropolis. *Journal of Animal and Veterinary Advances*, 8(5), 883–887.
- Witten, I. H., Frank, E., Hall, M. A., Pal, C. J., & DATA, M. (2005). Practical machine learning tools and techniques. *Data Mining*, 2(4).
- Yadav, A. K., Malik, H., & Chandel, S. S. (2014). Selection of most relevant input parameters using WEKA for artificial neural network based solar radiation prediction models. *Renewable and Sustainable Energy Reviews*, 31, 509–519.