

# Penerapan Algoritma Decision Tree dalam Klasifikasi Kualitas Penjualan Motor Bekas

Hengky<sup>1</sup>, Novianto<sup>2</sup>, Amalia Putri Yulandi<sup>3</sup>, Dandhytya Andrea Puspa<sup>4</sup>, Henly<sup>5</sup>

<sup>1-5</sup>Universitas Internasional Batam, email: [1931008.hengky@uib.ac.id](mailto:1931008.hengky@uib.ac.id), [1931037.novianto@uib.edu](mailto:1931037.novianto@uib.edu), [1931025.amalia@uib.edu](mailto:1931025.amalia@uib.edu), [1931044.dandhytya@uib.edu](mailto:1931044.dandhytya@uib.edu), [1931082.henly@uib.edu](mailto:1931082.henly@uib.edu).

## Abstrak:

Penjualan motor bekas kerap menjadi masalah dalam menentukan kualitas. Banyak yang belum mengetahui cara signifikan dalam menentukan kualitas penjualan motor bekas. Pada pelaku UMKM Liga Jaya Motor memiliki hambatan untuk mengetahui cara penentuan kualitas penjualan motor bekas. Untuk mengetahui penentuan kualitas penjualan motor bekas peneliti menggunakan pemodelan algoritma *Decision Tree* dengan aplikasi KNIME sebagai pembuatan *workflow*. Pemodelan *Decision Tree* dengan konsep klasifikasi kualitas penjualan motor bekas dilakukan dengan rasio 9:1 untuk digunakan dalam data latih dan data uji. Dari penelitian ini, sebagai pemilik sebelum nya yaitu owner 1<sup>st</sup> memiliki kualitas penjualan motor bekas yang paling signifikan berdasarkan dari hasil klasifikasi sehingga data ini dapat diterapkan pada UMKM Liga Jaya Motor. Algoritma *Decision Tree* cukup akurat untuk digunakan dalam klasifikasi kualitas penjualan motor bekas berdasarkan data Internet harga pasaran motor bekas.

Kata kunci : Kualitas penjualan motor bekas, *Decision Tree*, KNIME

## Abstract :

*The sale of used motorbikes is often a problem in determining quality. Many do not know a significant way in determining the quality of selling used motorbikes. The Liga Jaya Motor UMKM players have obstacles in knowing how to determine the quality of used motorbike sales. To determine the quality of used motorbike sales, researchers used the Decision Tree algorithm modeling with the KNIME application as a workflow creation. Decision Tree modeling with the concept of classifying the quality of used motorbike sales is carried out with a 9: 1 ratio for use in training data and test data. From this study, as the previous owner, the 1st owner has the most significant quality of used motorbike sales based on the classification results so that this data can be applied to the Liga Jaya Motor UMKM. The Decision Tree algorithm is accurate enough to be used in classifying the quality of used motorbike sales based on Internet data on used motorbike market prices.*

*Keywords : The quality of the sale of used motorbikes, Decision Tree, KNIME*

## 1. Pendahuluan

Machine learning adalah salah satu cabang dari AI yang dapat membuat sebuah komputer untuk berpikir dan belajar dengan sendirinya (Alzubi et al., 2018). Seperti manusia, mesin belajar dengan menggunakan data yang dibutuhkan untuk mempelajari suatu bidang. Dengan menggunakan ribuan bahkan hingga jutaan data tentang bidang tertentu, mesin dapat mempelajari dan mengubah tingkah lakunya dengan tujuan untuk memperbagus pekerjaannya hingga mendapatkan akurasi yang tinggi yang bertujuan untuk mendapatkan pilihan yang benar sesuai dengan data yang diberikan oleh pengguna terhadap mesin yang akan digunakan untuk mempelajarinya. Untuk mempelajari ilmu yang ingin dipelajari, mesin membutuhkan algoritma yang berfungsi sebagai arahan untuk belajar.

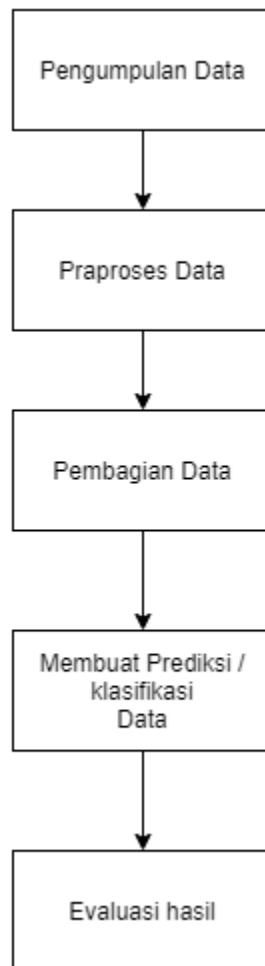
Salah satu algoritma yang digunakan pada machine learning adalah decision tree. Decision tree adalah pembelajaran yang diawasi oleh pengguna untuk mengawal proses pembelajaran pada mesin. Decision tree bekerja dengan membagi data berdasarkan parameter yang khusus sehingga data tersebut akan menjadi bercabang. Keunggulan dari algoritma ini adalah sangat cocok untuk

permasalahan regresi dan klasifikasi, sangat mudah dibaca, dan bagus untuk hasil yang berupa angka. Tetapi kekurangan dari algoritma ini adalah ketika parameter pada ilmu yang dipelajari sangat banyak, akan sangat sulit untuk membaca atau mengontrol besarnya decision tree tersebut (Ray, 2019).

Penentuan kualitas penjualan motor bekas menjadi salah satu permasalahan pada UMKM kami. Oleh karena itu, kami ingin mengetahui kualitas penjualan motor bekas dengan menggunakan algoritma decision tree untuk mempelajari data yang telah kami dapatkan dari internet yang berisi tentang harga pasaran motor second hand untuk menjadi pedoman pemilik UMKM kami dalam melakukan keputusan untuk membeli atau menjual motor second hand tersebut. Tujuan dari artikel ini adalah untuk membantu pemilik UMKM dalam melakukan keputusan dengan saran yang akan diberikan oleh algoritma decision tree tersebut.

## 2. Metodologi penelitian

Metodologi yang dilakukan pada penelitian ini dapat dilihat pada Gambar 1. Dimulai dari pengumpulan data penelitian dimana data tersebut akan digunakan sebagai bahan penelitian, kemudian data tersebut akan dilakukan praproses dimana pada tahapan ini data akan melakukan pembersihan sehingga data tersebut siap diolah. Sebelum masuk proses pembuatan prediksi / klasifikasi data, data akan dibagi menjadi data latih dan data uji. Data latih digunakan untuk pembentukan model, dan data uji digunakan untuk mengevaluasi hasil klasifikasi.



## Gambar 1. Alur Penelitian

### 2.1 Pengumpulan Data

Pada tahap ini adalah peneliti melakukan proses pengumpulan data yang akan digunakan sebagai bahan penelitian dalam proses pengolahan data. Data yang didapatkan dan digunakan oleh peneliti adalah data primer yang didapatkan dari situs dataset Kaggle dibuat oleh Nehal Birla sebagai bahan pembelajaran dalam melakukan penelitian data machine learning. Metode yang digunakan untuk mendapatkan data dengan melakukan riset di sebuah platform kumpulan data dan data tersebut disesuaikan dengan UMKM kami yaitu Liga Jaya motor yang bergerak di bidang perbengkelan.

### 2.2 Praproses Data

Pada tahap ini peneliti akan melakukan pembersihan data sebelum digunakan agar data siap diolah nantinya (Christian, 2020). Data yang dinyatakan tidak sempurna akan di buang. Sehingga hanya data yang terbaik bisa digunakan.

### 2.3 Pembagian Data

Pada tahap ini data akan dibagi menjadi rasio 9:1. Data yang digunakan untuk data latih sebesar 90% sedangkan untuk data uji mendapatkan 10%. Dengan begini diharapkan komputer dapat belajar lebih banyak dan memahami apa yang diminta.

### 2.4 Algoritma Decision Tree

Pada tahap ini peneliti menggunakan algoritma *Decision Tree*. Algoritma *Decision Tree* atau biasa disebut dengan pohon keputusan merupakan algoritma klasifikasi yang berbentuk seperti struktur pohon yang dapat menginterpretasikan pengetahuannya (Santoni et al., 2020). *Decision Tree* adalah salah satu metode klasifikasi yang paling digemari, dikarenakan mudah diimplementasikan oleh manusia. Selain itu, *decision tree* dapat digunakan sebagai mengeksplor data, menemukan hubungan tersembunyi antara sejumlah calon variabel yang akan diinput dengan sebuah variabel target. *Decision tree* juga dapat menggabungkan antara eksplorasi data dan pemodelan, sehingga sangat bagus sebagai Langkah awal dalam proses pemodelan bahkan ketika dijadikan sebagai model akhir dari beberapa teknik lain.

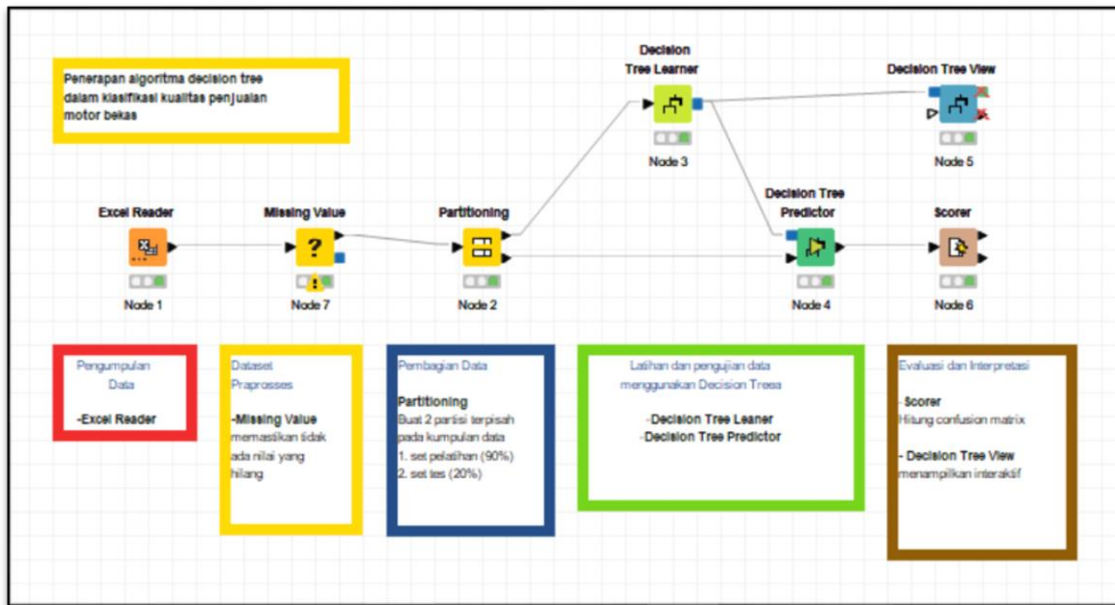
Lalu pengujian data yang digunakan oleh algoritma decision tree adalah pertama melalui *root node* dan terakhir adalah melalui *leaf node* yang akan menyimpulkan prediksi kelas bagi data tersebut. Pemilihan atribut untuk menjadi *rootnode* atau *internal node* sebagai atribut tes berdasarkan atas ukuran *impurity* dari masing-masing atribut. Ukuran-ukuran *impurity* yang umumnya digunakan adalah *information gain*, *gain ratio* dan *gini index*. Atribut yang memiliki nilai *impurity* tertinggi akan dipilih sebagai atribut tes (Kasih, 2019).

### 2.5 Evaluasi Hasil

Pada tahap ini peneliti melakukan evaluasi pada pola-pola tertentu yang dihasilkan oleh proses *Decision Tree Predictor* setelah melakukan data *testing* dengan memasukkan pemahaman algoritma yang sebelumnya telah dilatih. Evaluasi dilakukan dengan menerjemahkan pola-pola tersebut sehingga dapat lebih dipahami oleh pihak terkait.

## 3. Hasil pembahasan

Dengan menggunakan metodologi yang sudah jelaskan kami membuat *workflow* pada aplikasi KNIME ditunjukkan pada gambar dibawah ini:



**Gambar 2. Model penelitian menggunakan KNIME**

### 3.1 Pengumpulan Data

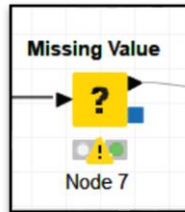
Dataset tersebut diperoleh dari situs internet *Kaggle* yang dibuat oleh Nihal Birla mengenai *Motorcycle Dataset* sesuai dengan UMKM yaitu Liga Jaya Motor yang bergerak dibidang perbengkelan dan mengenai motor. Dataset yang digunakan berjumlah 1061 data untuk menilai kualitas penjualan motor bekas dari pemilik sebelumnya. Di bawah ini adalah dataset yang kami miliki:

**Tabel 1. Dataset**

Row ID	S name	I selling_...	I year	S seller_t...	S owner	I km_driven	I ex_sho...
Row0	Royal Enfield Classic 350	175000	2019	Individual	1st owner	350	?
Row1	Honda Dio	45000	2017	Individual	1st owner	5650	?
Row2	Royal Enfield Classic G...	150000	2018	Individual	1st owner	12000	148114
Row3	Yamaha Fazer FI V 2.0...	65000	2015	Individual	1st owner	23000	89643
Row4	Yamaha SZ [2013-2014]	20000	2011	Individual	2nd owner	21000	?
Row5	Honda CB Twister	18000	2010	Individual	1st owner	60000	53857
Row6	Honda CB Hornet 160R	78500	2018	Individual	1st owner	17000	87719
Row7	Royal Enfield Bullet 35...	180000	2008	Individual	2nd owner	39000	?
Row8	Hero Honda CBZ extreme	30000	2010	Individual	1st owner	32000	?
Row9	Bajaj Discover 125	50000	2016	Individual	1st owner	42000	60122
Row10	Yamaha FZ16	35000	2015	Individual	1st owner	32000	78712
Row11	Honda Navi	28000	2016	Individual	2nd owner	10000	47255
Row12	Bajaj Avenger Street 220	80000	2018	Individual	1st owner	21178	95955
Row13	Yamaha YZF R3	365000	2019	Individual	1st owner	1127	351680
Row14	Jawa 42	185000	2020	Individual	1st owner	1700	?
Row15	Suzuki Access 125 [20...	25000	2012	Individual	1st owner	55000	58314
Row16	Hero Honda Glamour	25000	2006	Individual	1st owner	27000	?
Row17	Yamaha YZF R15 S	40000	2010	Individual	2nd owner	45000	117926
Row18	Royal Enfield Classic G...	150000	2018	Individual	1st owner	23000	148114
Row19	Yamaha FZ25	120000	2018	Individual	1st owner	39000	132680
Row20	Hero Passion Pro 110	15000	2008	Individual	1st owner	60000	?

### 3.2 Dataset Preprocesses

Pada tahap ini peneliti melakukan pemeriksaan data dan pembersihan data menggunakan node nilai yang hilang untuk memastikan tidak ada kesalahan dan kekosongan.



**Gambar 3. Praproses Data**

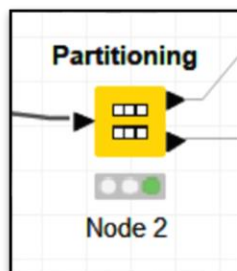
Berikut ini dataset yang telah dibersihkan :

**Tabel 2. Data Tabel dari node Missing Value**

Row ID	[S] name	[I] selling...	[I] year	[S] seller_t...	[S] owner	[I] km_driven	[I] ex_sho...
Row2	Royal Enfield Classic Gunmetal Grey	150000	2018	Individual	1st owner	12000	148114
Row3	Yamaha Fazer FI V 2.0 [2016-2018]	65000	2015	Individual	1st owner	23000	89643
Row5	Honda CB Twister	18000	2010	Individual	1st owner	60000	53857
Row6	Honda CB Hornet 160R	78500	2018	Individual	1st owner	17000	87719
Row9	Bajaj Discover 125	50000	2016	Individual	1st owner	42000	60122
Row10	Yamaha FZ16	35000	2015	Individual	1st owner	32000	78712
Row11	Honda Navi	28000	2016	Individual	2nd owner	10000	47255
Row12	Bajaj Avenger Street 220	80000	2018	Individual	1st owner	21178	95955
Row13	Yamaha YZF R3	365000	2019	Individual	1st owner	1127	351680
Row15	Suzuki Access 125 [2007-2016]	25000	2012	Individual	1st owner	55000	58314
Row17	Yamaha YZF R15 S	40000	2010	Individual	2nd owner	45000	117926
Row18	Royal Enfield Classic Gunmetal Grey	150000	2018	Individual	1st owner	23000	148114
Row19	Yamaha FZ25	120000	2018	Individual	1st owner	39000	132680
Row21	Honda Navi [2016-2017]	26000	2016	Individual	1st owner	17450	44389
Row22	Honda Activa i	32000	2013	Individual	2nd owner	20696	53900
Row25	Honda Dream Yuga	25000	2012	Individual	1st owner	35000	56147
Row27	Honda Navi [2016-2017]	42000	2017	Individual	1st owner	24000	44389
Row28	Yamaha Fazer [2009-2016]	40000	2013	Individual	3rd owner	35000	84751
Row30	Hero Glamour 125	55000	2017	Individual	1st owner	2500	61600
Row31	Yamaha FZ S [2012-2016]	38000	2013	Individual	1st owner	75000	79432
Row32	Hero Xtreme Sports	43000	2016	Individual	2nd owner	50000	78350
Row34	Honda CB Shine SP	62000	2019	Individual	1st owner	10000	71049
Row35	Honda Activa 5G	60000	2018	Individual	1st owner	6502	57557
Row36	Hero Glamour 125	45000	2018	Individual	1st owner	21000	61600
Row37	Honda CBR-250R	120000	2016	Individual	1st owner	16000	202310
Row41	Hero Glamour FI	50000	2018	Individual	1st owner	20000	71150
Row43	Hero Passion XPro	28000	2012	Individual	1st owner	81000	56750
Row44	Yamaha FZ S V 2.0	85000	2018	Individual	1st owner	8000	84042
Row45	Hero Achiever 150	29900	2016	Individual	1st owner	55000	69750

### 3.3 Pembagian Data

Pada tahap ini peneliti melakukan proses pemisahan dataset dengan rasio 9: 1 menggunakan node *partitioning* yang terdiri dari dataset latihan sebesar 90% dan dataset pengujian sebesar 10%. Dataset latihan akan dialirkan ke node selanjutnya untuk proses klasifikasi, sedangkan dataset pengujian digunakan sebagai dataset pada tahap pengujian.



**Gambar 4. Partitioning node**

Data yang digunakan untuk data Latihan :



**Tabel 3. Data Table yang digunakan untuk Data Training**

Table "default" - Rows: 563 Spec - Columns: 7 Properties Flow Variables							
Row ID	[S] name	[I] selling_...	[I] year	[S] seller_t...	[S] owner	[I] km_driven	[I] ex_sho...
Row2	Royal Enfield Classic Gunmetal Grey	150000	2018	Individual	1st owner	12000	148114
Row3	Yamaha Fazer FI V 2.0 [2016-2018]	65000	2015	Individual	1st owner	23000	89643
Row5	Honda CB Twister	18000	2010	Individual	1st owner	60000	53857
Row6	Honda CB Hornet 160R	78500	2018	Individual	1st owner	17000	87719
Row9	Bajaj Discover 125	50000	2016	Individual	1st owner	42000	60122
Row10	Yamaha FZ16	35000	2015	Individual	1st owner	32000	78712
Row11	Honda Navi	28000	2016	Individual	2nd owner	10000	47255
Row12	Bajaj Avenger Street 220	80000	2018	Individual	1st owner	21178	95955
Row13	Yamaha YZF R3	365000	2019	Individual	1st owner	1127	351680
Row15	Suzuki Access 125 [2007-2016]	25000	2012	Individual	1st owner	55000	58314
Row17	Yamaha YZF R15 S	40000	2010	Individual	2nd owner	45000	117926
Row18	Royal Enfield Classic Gunmetal Grey	150000	2018	Individual	1st owner	23000	148114
Row19	Yamaha FZ25	120000	2018	Individual	1st owner	39000	132680
Row21	Honda Navi [2016-2017]	26000	2016	Individual	1st owner	17450	44389
Row22	Honda Activa i	32000	2013	Individual	2nd owner	20696	53900
Row27	Honda Navi [2016-2017]	42000	2017	Individual	1st owner	24000	44389
Row28	Yamaha Fazer [2009-2016]	40000	2013	Individual	3rd owner	35000	84751
Row31	Yamaha FZ S [2012-2016]	38000	2013	Individual	1st owner	75000	79432
Row32	Hero Xtreme Sports	43000	2016	Individual	2nd owner	50000	78350
Row34	Honda CB Shine SP	62000	2019	Individual	1st owner	10000	71049
Row35	Honda Activa SG	60000	2018	Individual	1st owner	6502	57557
Row36	Hero Glamour 125	45000	2018	Individual	1st owner	21000	61600
Row37	Honda CBR-250R	120000	2016	Individual	1st owner	16000	202310
Row41	Hero Glamour FI	50000	2018	Individual	1st owner	20000	71150
Row43	Hero Passion XPro	28000	2012	Individual	1st owner	81000	56750
Row44	Yamaha FZ S V 2.0	85000	2018	Individual	1st owner	8000	84042

Berikut data yang digunakan untuk pengujian :

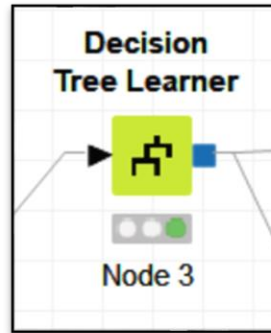
**Tabel 4. Data Table yang digunakan untuk Data Testing**

Table "default" - Rows: 63 Spec - Columns: 7 Properties Flow Variables							
Row ID	[S] name	[I] selling_...	[I] year	[S] seller_t...	[S] owner	[I] km_driven	[I] ex_sho...
Row25	Honda Dream Yuga	25000	2012	Individual	1st owner	35000	56147
Row30	Hero Glamour 125	55000	2017	Individual	1st owner	2500	61600
Row87	Royal Enfield Cla...	140000	2018	Individual	1st owner	22000	148114
Row114	Bajaj Discover 12...	17500	2013	Individual	1st owner	38000	58438
Row123	Honda CB Hornet...	55000	2017	Individual	1st owner	14000	87543
Row133	Bajaj Pulsar 180	15000	2008	Individual	1st owner	55000	85011
Row192	Bajaj Pulsar AS200	60000	2015	Individual	1st owner	50000	92074
Row202	Bajaj Pulsar 135 LS	56000	2017	Individual	3rd owner	37714	64589
Row210	Yamaha FZ S [20...	42000	2013	Individual	1st owner	60000	79432
Row227	TVS Scooty Zest ...	28000	2015	Individual	1st owner	32800	52372
Row250	Hero Maestro Edge	25000	2017	Individual	1st owner	55000	52790
Row282	Hero Maestro	25000	2015	Individual	2nd owner	40000	49412
Row299	Bajaj Avenger St...	40000	2016	Individual	1st owner	29689	80435
Row353	Suzuki Gixxer [20...	62000	2017	Individual	1st owner	15000	78316
Row359	Hero Pleasure	40000	2014	Individual	1st owner	14000	51600
Row360	Honda Aviator	41000	2013	Individual	1st owner	36500	63537
Row368	Hero Splendor Pl...	70000	2019	Individual	1st owner	3000	55600
Row370	Hero Maestro	26500	2016	Individual	2nd owner	40000	53830
Row376	Yamaha FZ25	130000	2017	Individual	1st owner	4000	132680
Row386	Yamaha FZ V 2.0	35000	2014	Individual	1st owner	13500	81508
Row414	Hero Splendor iS...	40000	2015	Individual	1st owner	19055	52271
Row417	Honda Aviator	28000	2013	Individual	1st owner	28087	63537
Row442	Honda CBR-250R	100000	2011	Individual	2nd owner	3500	171646
Row449	Royal Enfield Cla...	195000	2018	Individual	1st owner	2845	164095
Row457	Mahindra Flyte	22000	2009	Individual	3rd owner	20000	44493

### 3.4 Latihan dan pengujian data dengan Decision Tree Algoritma

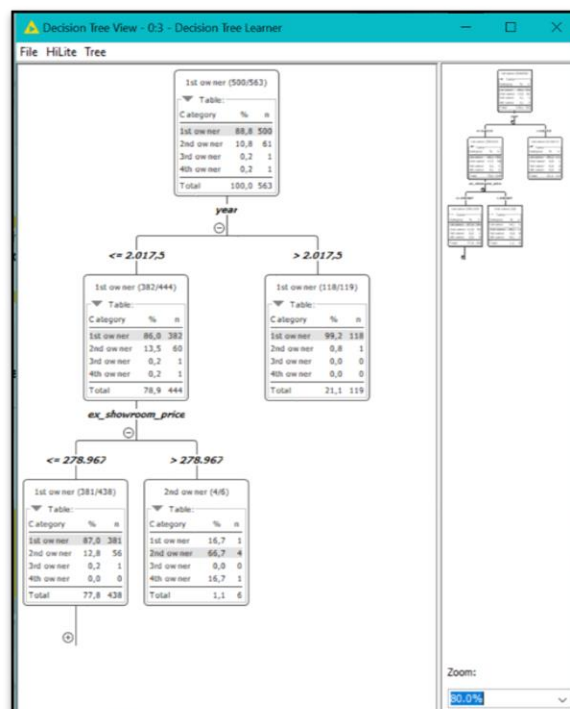
Setelah melakukan pembagian data, data akan dibagikan sesuai dengan yang diminta. Data pertama sebesar 90% dari jumlah data akan digunakan sebagai data Latihan atau disebut dengan data *training* dilakukan oleh *Decision Tree Learner Node*. Sedangkan 10% dari jumlah data akan

digunakan sebagai data pengujian atau yang disebut data *testing* dilakukan oleh *Decision Tree Predictor Node* dengan pengetahuan yang dipelajari pada tahap *training*.

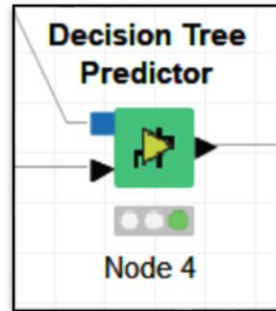


**Gambar 5. Decision Tree Learner Node**

Tahap selanjutnya adalah klasifikasi dari data *training* menggunakan algoritma *Decision Tree Learner*. Hasil yang didapat pada tahap ini adalah dataset yang telah di klasifikasi sesuai dengan bagan *Decision Tree*.

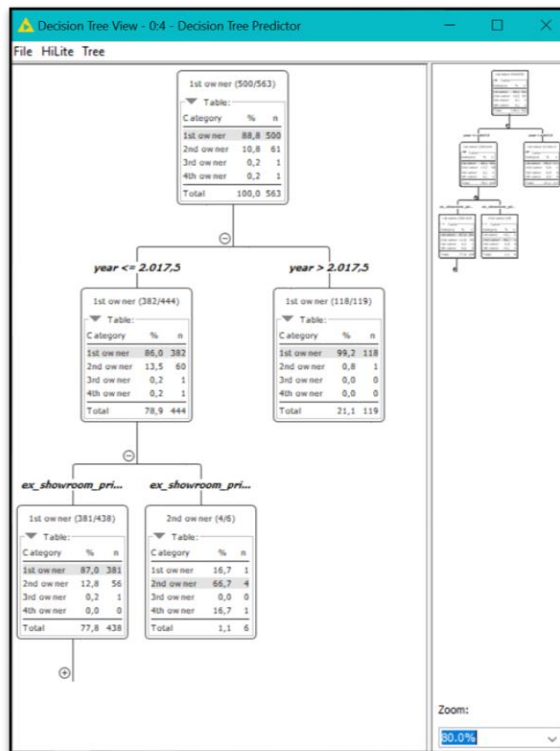


**Gambar 6. Hasil dari klasifikasi Decision Tree Learner**



**Gambar 7. Decision Tree Predictor Node**

Tahap selanjutnya adalah klasifikasi dari data *testing* menggunakan algoritma *Decision Tree Predictor*. Hasil yang didapat pada tahap ini adalah dataset yang telah di klasifikasi sesuai dengan bagan *Decision Tree* yang diinginkan peneliti.

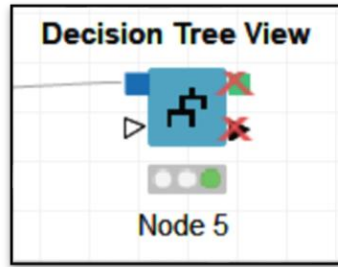


**Gambar 8. Hasil dari klasifikasi Decision Tree Predictor**

### 3.5 Evaluasi dan Interpretasi

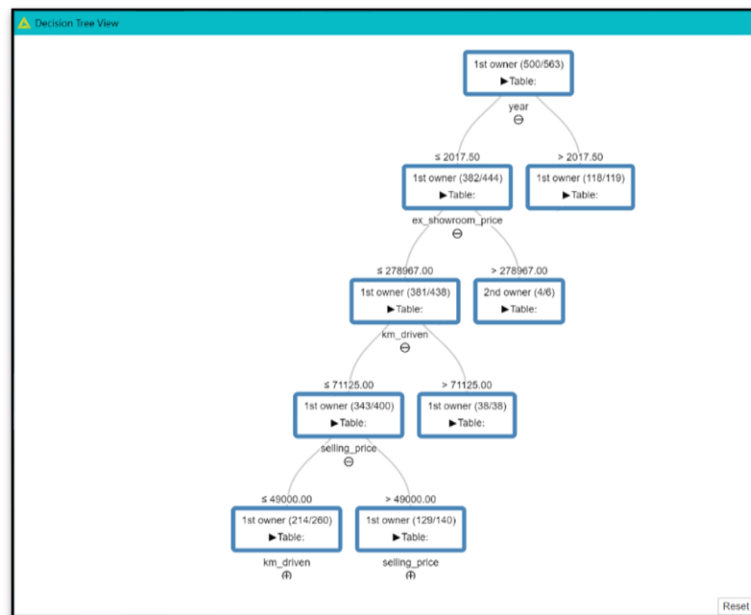
Pada tahap ini data klasifikasi akan dialirkan ke dua node. Node pertama digunakan untuk menampilkan Pohon Keputusan interaktif.





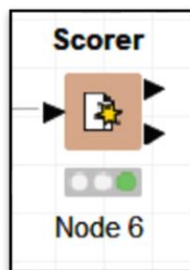
**Gambar 9. Decision Tree View Node**

Berikut contoh tampilan *Decision Tree* dalam bentuk interaktif dari hasil klasifikasi :



**Gambar 10. Hasil Decision Tree Interactive**

Node kedua digunakan untuk menyimpulkan data klasifikasi sebagai bahan perbandingan sekaligus menghasilkan nilai akhir dari klasifikasi yang sudah dibuat. Data yang dihasilkan pada node ini akan digunakan untuk menarik kesimpulan akhir dari dataset Latihan dengan dataset yang diuji dari tahap awal hingga akhir. Untuk menentukan nilai akhir menggunakan *Node Scorer*.



**Gambar 11. Scorer Node**

Data dibawah ini merupakan data yang dihasilkan sebagai kesimpulan dari klasifikasi yang telah dibuat. Untuk nilai *F1 Score* kami menggunakan rumus  $F1\ Score = 2 * (Recall * Precision) / (Recall + Precision)$  (Basha et al., 2019). Sehingga hasil yang kami dapatkan:

$$F1\ Score = 2 * (0.875 * 0.875) / (0.875 + 0.875)$$

$$F1\ Score = 2 * (0.765625) / (1.75)$$

$$F1\ Score = 2 * 0.4375$$

$$F1\ Score = 0.875$$

Accuracy statistics - 0:6 - Scorer												
File Edit Hilite Navigation View												
Table "default" - Rows: 5 Spec - Columns: 11 Properties Flow Variables												
Row ID	I TruePo...	I FalsePo...	I TrueNe...	I FalseN...	D Recall	D Precision	D Sensitivity	D Specifcity	D F-meas...	D Accuracy	D Cohen...	
1st owner	49	7	0	7	0.875	0.875	0.875	0	0.875	?	?	
2nd owner	0	7	51	5	0	0	0	0.879	NaN	?	?	
3rd owner	0	0	61	2	0	?	0	1	?	?	?	
4th owner	0	0	63	0	?	?	?	1	?	?	?	
Overall	?	?	?	?	?	?	?	?	?	0.778	-0.105	

**Gambar 12. Hasil Accuracy Statistics – 0:6 – Scorer**

Confusion Matrix - 0:6 - Scorer					
File Hilite					
owner \ Pr...	1st owner	2nd owner	3rd owner	4th owner	
1st owner	49	7	0	0	
2nd owner	5	0	0	0	
3rd owner	2	0	0	0	
4th owner	0	0	0	0	
Correct classified: 49 Wrong classified: 14					
Accuracy: 77,778 % Error: 22,222 %					
Cohen's kappa (κ) -0,105					

**Gambar 13. Hasil Confusion Matrix – 0:6 – Scorer**

Sehingga parameter yang dihasilkan pada data diatas dengan dataset yang kami gunakan memiliki nilai : *F1 Score* : 87.5%, *Recall* : 87.5%, *Precision* : 87.5%, dan *Accuracy* : 77.778%.

#### 4. Kesimpulan

Penerapan machine learning dengan menggunakan algoritma *Decision Tree* telah kami gunakan untuk mengetahui hasil penelitian dari kualitas penjualan motor bekas. Adapun metodologi yang kami gunakan pada aplikasi KNIME, sehingga hasil penelitian bisa menjadi bahan pembahasan pada permasalahan UMKM kami.

Hasil penelitian ini mendapatkan nilai *F1 Score* : 87.5%, *Recall* : 87.5%, *Precision* : 87.5%, dan *Accuracy* : 77.778%. Dan dari klasifikasi *Decision Tree* penjualan motor bekas menyatakan bahwa pemilik sebelumnya yaitu owner 1st memiliki signifikasi yang baik sebagai kualitas penjualan motor bekas. Dapat diketahui dari hasil penelitian, disimpulkan bahwa algoritma

*Decision Tree* adalah algoritma yang baik digunakan untuk menilai data yang digunakan pada penelitian ini, yaitu menilai data dalam kualitas penjualan motor bekas.

#### **Daftar Pustaka**

- Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. *Journal of Physics: Conference Series*, 1142(1). <https://doi.org/10.1088/1742-6596/1142/1/012012>
- Basha, S. M., Bagyalakshmi, K., Ramesh, C., Rahim, R., Manikandan, R., & Kumar, A. (2019). Comparative study on performance of document classification using supervised machine learning algorithms: KNIME. *International Journal on Emerging Technologies*, 10(1), 148–153.
- Christian, Y. (2020). Application of K-Means Algorithm for Clustering the Quality of Lecturer Learning at Batam International University. *International Journal of Information System & Technology*, 3(36), 191–199.
- Kasih, P. (2019). Pemodelan Data Mining Decision Tree Dengan Classification Error Untuk Seleksi Calon Anggota Tim Paduan Suara. *Innovation in Research of Informatics (INNOVATICS)*, 1(2), 63–69. <https://doi.org/10.37058/innovatics.v1i2.918>
- Ray, S. (2019). A Quick Review of Machine Learning Algorithms. *Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing: Trends, Prespectives and Prospects, COMITCon 2019*, 35–39. <https://doi.org/10.1109/COMITCon.2019.8862451>
- Santoni, M. M., Chamidah, N., & Matondang, N. (2020). Prediksi Hipertensi menggunakan Decision Tree, Naïve Bayes dan Artificial Neural Network pada software KNIME. *Techno.Com*, 19(4), 353–363. <https://doi.org/10.33633/tc.v19i4.3872>