

Tugas Project

LAPORAN PENGANTAR DATA MINING
PENGAPLIKASIAN DECISION TREE DAN ASSOCIATION RULES PADA
DATA HOUSE PRICE PREDICTION DAN GROCERY SHOP



DISUSUN OLEH :

Marhama / H051191029

Andi Apridhani Mattalatta / H051191031

Isal Sulkarnain / H051191036

DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN

2022

Kata Pengantar

Assalamualaikum Warahmatullahi Wabarahkatuh, Pertama-tama marilah kita panjatkan puji syukur kita atas kehadiran Allah SWT. karena berkat limpahan rahmat dan taufiq hidayahnya sehingga kami dapat menyelesaikan Laporan Pengantar Data Mining yang berjudul “Pengaplikasian Decition Tree Dan Association Rules Pada Data House Price Prediction Dan Grocery Shop”. Kami berterima kasih kepada dosen kami yang telah membimbing kami sehingga kami dapat menyelesaikan laporan ini dengan baik.

Laporan ini telah kami susun dengan maksimal dan mendapatkan bantuan dari berbagai pihak sehingga dapat memperlancar pembuatan laporan ini. Untuk itu kami menyampaikan banyak terima kasih kepada semua pihak yang telah berkontribusi dalam pembuatan laporan ini.

Terlepas dari semua itu, kami menyadari sepenuhnya bahwa masih ada kekurangan baik dari segi susunan kalimat maupun tata bahasanya. Oleh karena itu dengan tangan terbuka kami menerima segala saran dan kritik dari pembaca agar kami dapat memperbaiki laporan ini.

Akhir kata kami berharap semoga laporan yang berjudul “Pengaplikasian Decition Tree Dan Association Rules Pada Data House Price Prediction Dan Grocery Shop” ini dapat memberikan manfaat maupun inspirasi terhadap pembaca.

Makassar, 8 April 2022

Penulis

DAFTAR ISI

Kata Pengantar	2
DAFTAR ISI.....	3
BAB 1 PENDAHULUAN	4
A. Latar Belakang	4
B. Rumusan Masalah.....	4
C. Tujuan	4
BAB 2 TINJAUAN PUSTAKA	5
A. Data Mining	5
B. Pohon keputusan (Decision tree)	6
C. Analisis asosiasi (Association tree)	7
BAB 3 METODOLOGI PENELITIAN	7
A. Sumber data	8
B. Teknik analisis data.....	8
BAB 4 ANALISIS DATA	8
A. Decision Tree dan Evaluasi Kinerja Pengklasifikasian	9
B. Association Rules.....	14
BAB 5 KESIMPULAN	18

BAB 1

PENDAHULUAN

A. Latar Belakang

Seiring perkembangan teknologi dan komunikasi, maka semakin berkembang pula bisnis diberbagai bidang, contohnya dibidang property dan perdagangan. Sudah tidak diragukan lagi bahwa bisnis di bidang properti memiliki prospek yang menarik untuk jangka panjang. Apalagi perkembangan dari penduduk saat ini terus mengalami pertumbuhan yang sangat pesat. Sehingga kebutuhan akan tempat tinggal pun semakin bertambah dan membuat bisnis properti tidak akan mati kalau dijalankan dengan sabar dan tekun. Meskipun bisnis ini harus dimulai dengan modal yang cukup besar, tapi keuntungan yang didapatkan bisa jauh lebih besar dari modal awal yang dikeluarkan. Sehingga dalam melakukan analisis pada kasus tersebut lebih tepat menggunakan *Decision Tree*.

Pada bisnis perdagangan, diperlukan strategi dalam menempatkan barang. Hal ini sangat berpengaruh pada tingkat penjualan nantinya. Dalam suatu hari penjualan bisa terjadi banyak transaksi, terutama pada proses distribusi penjualan . Karena jumlahnya yang sangat banyak maka data sulit untuk dianalisis. Dari data yang sangat banyak tersebut toko ingin adanya pengolahan data lebih lanjut untuk menemukan informasi barang apa yang dibeli secara bersamaan oleh pelanggan sebagai penjual bisa mengambil keputusan dalam menempatkan suatu barang secara berdekatan. Maka untuk mengetahui hal tersebut, dilakukan analisis menggunakan *Assosiacion Rules*.

B. Rumusan Masalah

Adapun rumusan masalah pada laporan ini sebagai berikut :

1. Buatlah decision tree dari dataset pada bidang property
2. Buatlah association rules dari dataset pada bidang perdagangan

C. Tujuan

Adapun tujuan pembuatan makalah ini sebagai berikut :

1. Buatlah decision tree dari dataset pada bidang property
2. Buatlah association rules dari dataset pada bidang perdagangan

BAB 2

TINJAUAN PUSTAKA

A. Data Mining

Menurut Mardi (2016) *Data mining* adalah proses yang menggunakan 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* merupakan serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual.

Menurut Pramadhani dan Setiadi (2014) Secara sederhana, data mining atau penambangan data dapat didefinisikan sebagai proses seleksi, eksplorasi, dan pemodelan dari sejumlah besar data untuk menemukan pola atau kecenderungan yang biasanya tidak disadari keberadaannya. *Data mining* dapat dikatakan sebagai proses mengekstrak pengetahuan dari sejumlah besar data yang tersedia. Pengetahuan yang dihasilkan dari proses *data mining* harus baru, mudah dimengerti, dan bermanfaat. Dalam *data mining*, data disimpan secara elektronik dan diproses secara otomatis oleh komputer menggunakan teknik dan perhitungan tertentu.

Sebagai suatu rangkaian proses, data mining dapat dibagi menjadi beberapa tahap. Tahap-tahap tersebut bersifat interaktif di mana pemakai terlibat langsung atau dengan perantaraan *knowledge base*, berikut tahapan-tahapannya (Tampubolon & Saragi 2013) :

- a. Pembersihan Data (*Data Cleaning*) : Pembersihan data digunakan untuk menghilangkan noise dan data yang tidak konsisten atau data tidak relevan.
- b. Integrasi Data (*Data Integration*) : Integrasi data merupakan penggabungan data dari berbagai basis data ke dalam satu basis data baru (warehouse). Karena dengan data warehouse, data digabungkan dengan struktur yang efisien.
- c. Seleksi Data (*Data Selection*) : Data yang ada pada basis data sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari basis data.
- d. Transformasi Data (*Data Transformation*) : Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining.

- e. Proses Mining : Merupakan penerapan suatu metode untuk menemukan *rule* atau aturan dari data.
- f. Evaluasi Pola (*Pattern Evaluation*) : Dalam tahap ini hasil dari teknik data mining berupa rule/aturan dievaluasi untuk mengetahui apakah hipotesa yang ada memang tercapai. Bila hasil yang diperoleh tidak sesuai hipotesa, ada beberapa alternatif yang dapat diambil seperti memperbaiki proses data mining, mencoba metode *data mining* lain, atau menerima hasil ini sebagai suatu hasil yang mungkin bermanfaat.
- g. Presentasi Pengetahuan (*Knowledge Presentation*) : Merupakan visualisasi dari aturan mengenai metode yang digunakan untuk memudahkan pengguna memperoleh informasi.

B. Pohon keputusan (Decision tree)

Menurut Haryati, Sudarsono dan Suryana (2015) Pohon keputusan adalah salah satu metode klasifikasi yang paling populer karena mudah diinterpretasi manusia. Pohon keputusan adalah model prediksi menggunakan struktur pohon atau struktur berhirarki. Konsep dari pohon keputusan adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan. Data dalam pohon keputusan biasanya dinyatakan dalam bentuk tabel dengan atribut dan *record*. Atribut menyatakan suatu parameter yang dibuat sebagai kriteria dalam pembentukan tree. Misalkan untuk menentukan main tenis, kriteria yang digunakan adalah cuaca, angin, iklim dan temperatur.

Manfaat utama menggunakan pohon keputusan adalah kemampuannya untuk membreak down proses pengambilan keputusan yang kompleks menjadi lebih simpel sehingga pengambilan keputusan akan menjadi lebih menginterpretasikan solusi permasalahan. Pohon keputusan juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target. Pohon keputusan memadukan antara eksplorasi data dan pemodelan sehingga sangat bagus sebagai langkah awal pemodelan bahkan ketika dijadikan sebagai model akhir dari beberapa teknik lain.

Menurut Ardi Ramadhan Sukma, Riqadri Halfis dan Ady Hermawan (2019) Pohon keputusan adalah sebuah struktur yang dapat digunakan untuk mengubah data menjadi pohon keputusan yang akan menghasilkan aturan-aturan keputusan besar

menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan. Pohon keputusan adalah salah satu teknik penambangan data yang paling populer untuk penemuan pengetahuan. Secara sistematis menganalisis dan mengekstrak aturan untuk tujuan Klasifikasi / prediksi.

C. Analisis asosiasi (Association tree)

Analisis asosiasi atau *association rule* mining adalah teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item (Kusrini, dkk, 2009). Algoritma aturan asosiasi akan menggunakan data latihan, sesuai dengan pengertian data mining untuk menghasilkan pengetahuan. Aturan asosiasi yang berbentuk “*if...then...*” atau “*jika...maka...*” merupakan pengetahuan yang dihasilkan dari fungsi aturan asosiasi. Contoh asosiasi yang biasa terjadi seperti seberapa besar kemungkinan pembeli membeli roti dengan selai secara bersamaan. Penting tidaknya suatu aturan asosiatif dapat diketahui dengan dua parameter yaitu *support* dan *confidence*.

1. *Support* (nilai penunjang) adalah presentase kombinasi item tersebut dalam databases.
2. *Confidence* (nilai kepastian) adalah kuatnya hubungan antar item dalam aturan asosiasi.

Analisis asosiasi didefinisikan sebagai suatu proses untuk menemukan semua aturan asosiasi yang memenuhi syarat minimum untuk support (*minimum support*) dan syarat minimum untuk confidence (*minimum confidence*). Jika support-nya \geq minimum support dan confidence-nya \geq *minimum confidence*, maka rule tersebut bisa dikatakan sebagai *interesting rule*

BAB 3

METODOLOGI PENELITIAN

A. Sumber data

Data yang diolah merupakan data sekunder yang diperoleh di website kaggle. Adapun jenis data yang digunakan adalah data House Price Prediction dan Grocery Shop. Berikut link akses nya : <https://www.kaggle.com/datasets/shazadudwadia/supermarket> dan <https://www.kaggle.com/datasets/shree1992/housedata?resource=download>

B. Teknik analisis data

Teknik analisis data yang digunakan adalah Decision Tree untuk data House Prediction dan Association Rules untuk data Grocery Shop. Decision tree adalah diagram yang bisa membantumu memilih salah satu dari beberapa pilihan tindakan. Sedangkan association rules adalah salah satu task data mining deskriptif yang bertujuan untuk menemukan aturan asosiatif antara item-item data.

BAB 4

ANALISIS DATA

A. Decision Tree dan Evaluasi Kinerja Pengklasifikasian

1. Analisis data

Menginput data dan dilanjutkan dengan membuat data frame menggunakan fungsi data.frame

```
> input <- read.csv(file="E:\\Kuliah\\Semester 6\\Pengantar Data Mining\\archive-house  
prediction\\data.csv")  
> data_HousePrediction <- data.frame(input)  
> data_HousePrediction
```

Mengetahui jenis data dengan menggunakan fungsi str dan dimensi data dengan menggunakan fungsi dim

```
>str(data_HousePrediction)  
>View(data_HousePrediction)  
>dim(data_HousePrediction)
```

Membagi data mentah menjadi data train dan data test. Sebelumnya dilakukan pengacakan dengan fungsi set.seed. dalam membagi data train dan data test digunakan perbandingan 80% dan 20%.

```
set.seed(123456)  
sample_HousePrediction <- sample(2, nrow(data_HousePrediction), replace=TRUE,  
prob=c(0.8, 0.2))  
train_HP<-data_HousePrediction[sample_HousePrediction ==1,]  
test_HP<-data_HousePrediction[sample_HousePrediction ==2,]  
dim(train_HP)  
dim(test_HP)  
View(train_HP)  
View(test_HP)
```

Membuat formula prediksi menggunakan library (party)

```
> library(party)  
> myFormula <- price ~ bedrooms + bathrooms + sqft_basement
```

```
> HP_ctree <- ctree(myFormula, data=train_HP)
```

Melakukan pengecekan prediksi dengan fungsi predict

```
> table(predict(HP_ctree), train_HP$price)
```

Membuat plot decision tree dengan type simple

```
print(HP_ctree)
plot(HP_ctree)
plot(HP_ctree, type="simple")
```

Menghitung akurasi prediksi, akurasi dari prediksi ini menggunakan data test

```
Prediksi1 <- predict(HP_ctree, data=test_HP[, -18])
Hasilprediksi <- table(test_HP[, 18], predict(HP_ctree, test_HP[, -18]))
Hasilprediksi
accuracy <- (sum(diag(Hasilprediksi)))/sum(Hasilprediksi)
accuracy
```

Evaluasi kinerja menggunakan library (rpart)

```
library(rpart)
myFormula <- price ~ bedrooms + bathrooms + sqft_basement
HP_rpart <- rpart(myFormula, data = train_HP, control = rpart.control(minsplit = 10))
attributes(HP_rpart)
```

```
plot(HP_rpart)
text(HP_rpart, use.n=T)
```

```
opt <- which.min(HP_rpart$cptable[, "xerror"])
cp <- HP_rpart$cptable[opt, "CP"]
HP_prune <- prune(HP_rpart, cp = cp)
print(HP_prune)
plot(HP_prune)
```

```
text(HP_prune, use.n=T)
```

```
Price_pred<-predict(HP_prune, newdata = test_HP)
```

```
xlim<-range(data_HousePrediction$price)
```

```
plot(Price_pred~price,data = test_HP,xlab = "Observed",ylab = "Predicted", ylim=xlim,
```

```
xlim=xlim)
```

```
abline(a=0, b=1)
```

2. Interpretasi dari output

Untuk data_HousePrediction

	date	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront	view	condition	sqft_above	sqft_basement	yr_built	yr_renovated	street	city	statezip
1	2014-05-02 00:00:00	313000	3	1.50	1340	7912	1.5		0	0	3	1340	0	1955	2005	18810 Densmore	
2	2014-05-02 00:00:00	2384000	5	2.50	3650	9050	2.0	0	4	5	3370	280	1921	0	709 W Blaine St		
3	2014-05-02 00:00:00	342000	3	2.00	1930	11947	1.0	0	0	4	1930	0	1966	0	26206-26214 14:		
4	2014-05-02 00:00:00	420000	3	2.25	2000	8030	1.0	0	0	4	1000	1000	1963	0	857 170th Pl NE		
5	2014-05-02 00:00:00	550000	4	2.50	1940	10500	1.0	0	0	4	1140	800	1976	1992	9105 170th Ave f		
6	2014-05-02 00:00:00	490000	2	1.00	880	6380	1.0	0	0	3	880	0	1938	1994	522 NE 88th St		
7	2014-05-02 00:00:00	335000	2	2.00	1350	2560	1.0	0	0	3	1350	0	1976	0	2616 174th Ave f		
8	2014-05-02 00:00:00	462000	4	2.50	2710	35868	2.0	0	0	3	2710	0	1989	0	23762 SE 253rd f		
9	2014-05-02 00:00:00	452500	3	2.50	2430	88426	1.0	0	0	4	1570	860	1985	0	46611-46625 SE		
10	2014-05-02 00:00:00	640000	4	2.00	1520	6200	1.5	0	0	3	1520	0	1945	2010	6811 55th Ave N		

Showing 1 to 11 of 4,600 entries, 18 total columns

Untuk train_HP

	date	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront	view	condition	sqft_above	sqft_basement	yr_built	yr_renovated	street	city	statezip
1	2014-05-02 00:00:00	313000	3	1.50	1340	7912	1.5		0	0	3	1340	0	1955	2005	18810 Densmore	
2	2014-05-02 00:00:00	2384000	5	2.50	3650	9050	2.0	0	4	5	3370	280	1921	0	709 W Blaine St		
3	2014-05-02 00:00:00	342000	3	2.00	1930	11947	1.0	0	0	4	1930	0	1966	0	26206-26214 14:		
4	2014-05-02 00:00:00	420000	3	2.25	2000	8030	1.0	0	0	4	1000	1000	1963	0	857 170th Pl NE		
6	2014-05-02 00:00:00	490000	2	1.00	880	6380	1.0	0	0	3	880	0	1938	1994	522 NE 88th St		
7	2014-05-02 00:00:00	335000	2	2.00	1350	2560	1.0	0	0	3	1350	0	1976	0	2616 174th Ave f		
8	2014-05-02 00:00:00	462000	4	2.50	2710	35868	2.0	0	0	3	2710	0	1989	0	23762 SE 253rd f		
9	2014-05-02 00:00:00	452500	3	2.50	2430	88426	1.0	0	0	4	1570	860	1985	0	46611-46625 SE		
10	2014-05-02 00:00:00	640000	4	2.00	1520	6200	1.5	0	0	3	1520	0	1945	2010	6811 55th Ave N		
11	2014-05-02 00:00:00	463000	3	1.75	1710	7320	1.0	0	0	3	1710	0	1948	1994	Burke-Gilman Tr		

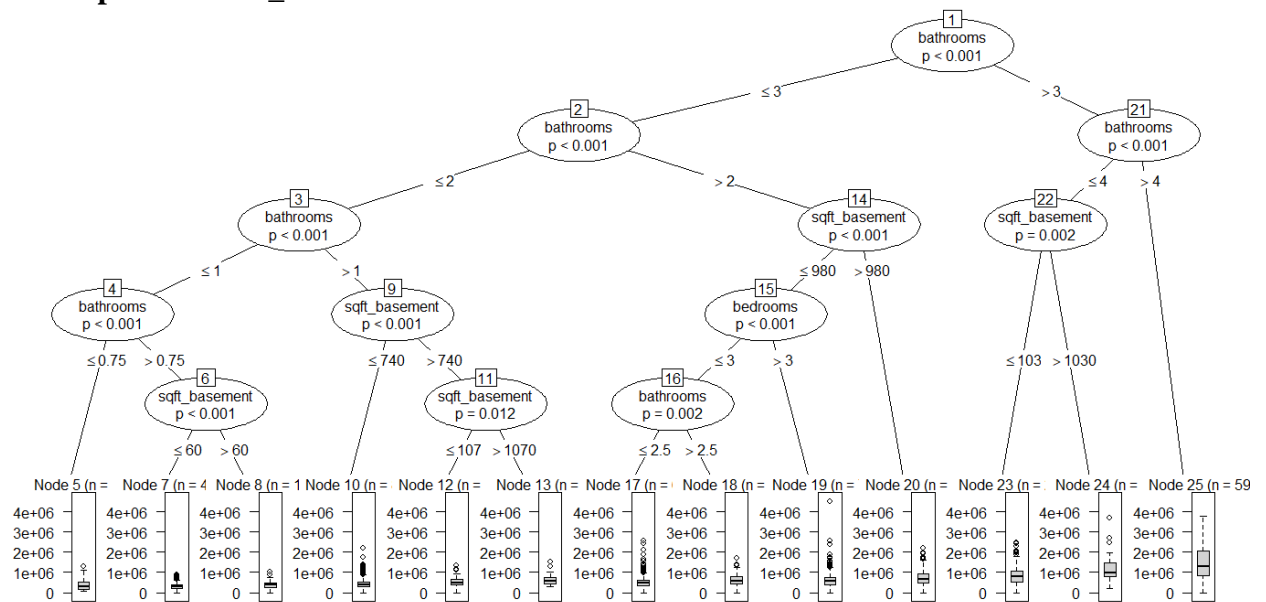
Showing 1 to 11 of 3,629 entries, 18 total columns

Untuk test_HP

	date	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront	view	condition	sqft_above	sqft_baseament	yr_built	yr_renovated	street
5	2014-05-02 00:00:00	550000	4	2.50	1940	10500	1.0		0	0	4	1140	800	1976	1992 9105 170th Ave I
14	2014-05-02 00:00:00	365000	3	1.00	1090	6435	1.0		0	0	4	1090	0	1955	2009 2504 SW Portlan
16	2014-05-02 00:00:00	242500	3	1.50	1200	9720	1.0		0	0	4	1200	0	1965	0 14034 SE 201st S
26	2014-05-02 00:00:00	285000	3	2.50	2090	10834	1.0		0	0	4	1360	730	1987	0 27736 23rd Aven
28	2014-05-02 00:00:00	698000	4	2.25	2200	11250	1.5		0	0	5	1300	900	1920	0 1036 4th St
29	2014-05-02 00:00:00	675000	5	2.50	2820	67518	2.0		0	0	3	2820	0	1979	2014 23525 SE 32nd V
36	2014-05-02 00:00:00	604000	3	2.50	3240	33151	2.0		0	2	3	3240	0	1995	0 30822 36th Ct Sv
39	2014-05-02 00:00:00	403000	3	2.00	1960	13100	1.0		0	2	5	1650	310	1957	0 17825 4th Ave Sl
40	2014-05-02 00:00:00	750000	3	2.50	2390	6550	1.0		0	2	4	1440	950	1955	2009 3628 59th Ave Sl
50	2014-05-02 00:00:00	838000	4	2.50	3310	42998	2.0		0	0	3	3310	0	2001	0 15712 NE 136th

Showing 1 to 11 of 971 entries, 16 total columns

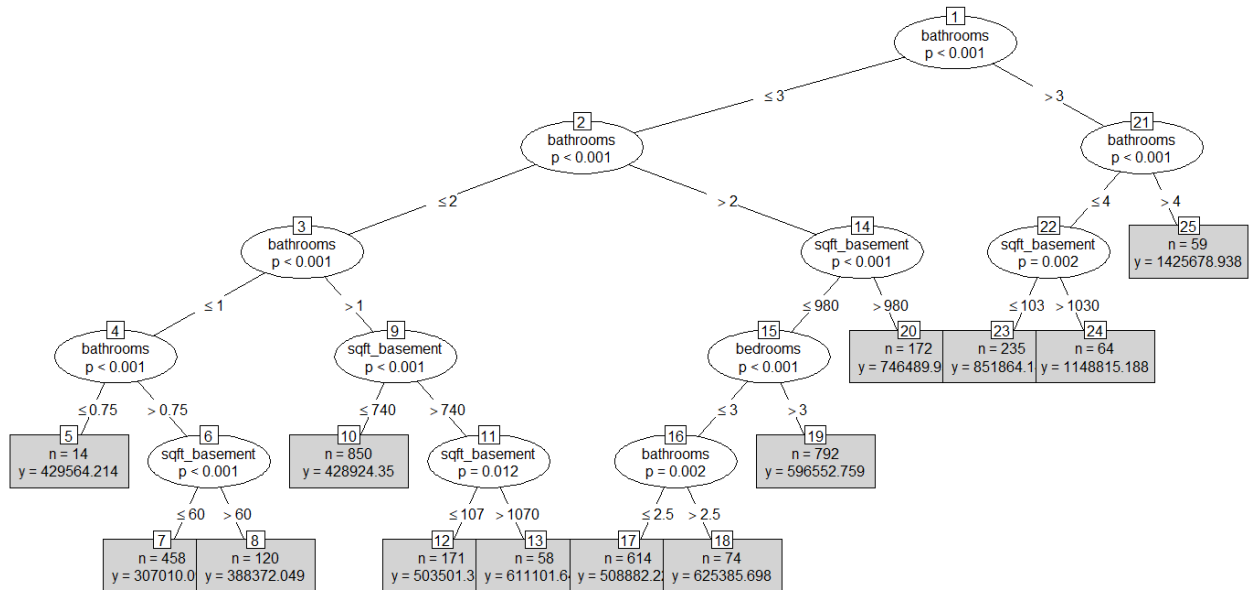
Untuk plot dari HP_ctree



Interpretasi :

Berdasarkan pohon klasifikasi diatas, didapatkan 6 pohon keputusan. Sebuah rumah yang memiliki kamar mandi lebih dari 3 dan mempunyai basement dengan luas lebih dari 1030 cenderung lebih mahal dengan total rumah sebanyak 59 rumah. Kemudian rumah yang tidak memiliki kamar mandi dan memiliki luas basement kurang dari 60 cenderung lebih murah dengan total rumah sebanyak 458 rumah.

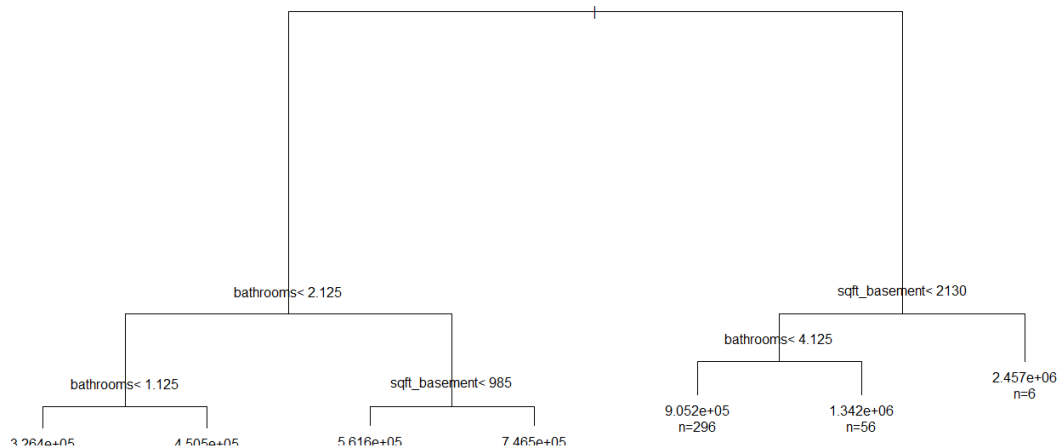
Untuk plot data HP_ctree dengan menggunakan type="simple"



Interpretasi :

Berdasarkan pohon klasifikasi diatas, didapatkan 6 pohon keputusan. Sebuah rumah yang memiliki kamar mandi lebih dari 4 cenderung lebih mahal dengan total rumah sebanyak 59 dan harga rumah 1.425.678,938. Kemudian rumah yang tidak memiliki kamar mandi dan memiliki luas basement kurang dari 60 cenderung lebih murah dengan total rumah sebanyak 458 rumah dan harga rumah 307.010.

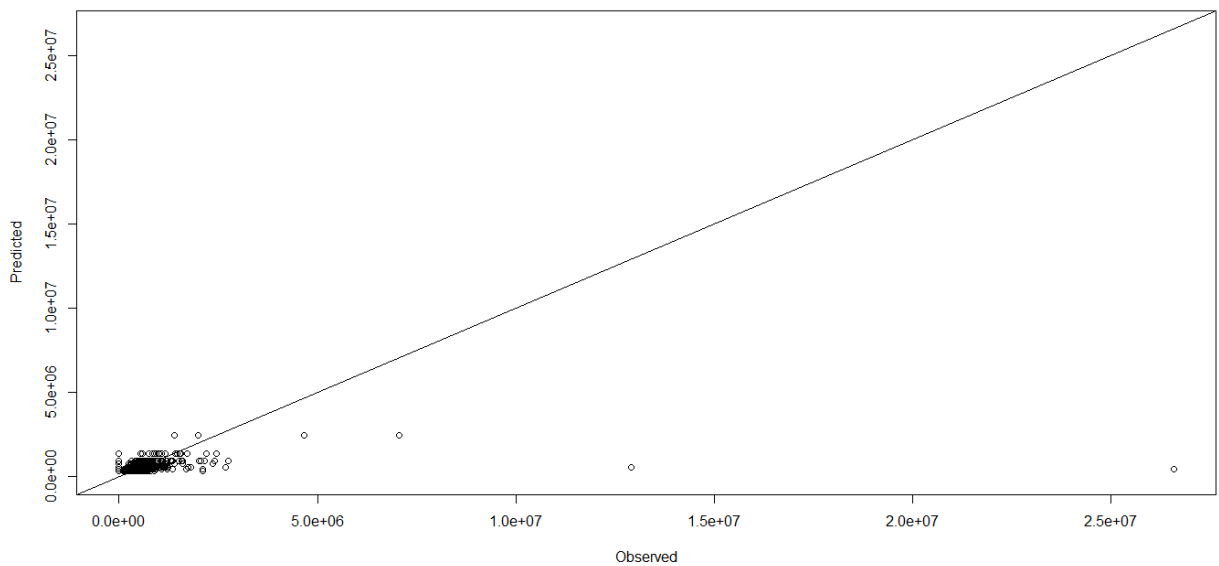
Untuk plot data HP_rpart



Interpretasi :

Dari plot yang diperoleh diatas dapat disimpulkan bahwa kamar tidur yang lebih banyak dari 3, dan memiliki basement dengan luas lebih dari 2130, dan tidak mempunyai kamar mandi maka harga yang ditawarkan untuk rumah tersebut sebesar 2.457.000.000. Untuk rumah yang mempunyai kamar tidur kurang dari 3, kamar mandi dengan luas kurang dari 1.125 ditawarkan dengan harga 326.400.000, sedangkan untuk rumah yang mempunyai kamar tidur kurang dari 3, kamar mandi dengan luas lebih dari 2.125, basement dengan luas lebih dari 985 seharga 746.500.000

Untuk abline



Interpretasi :

Dari plot data dapat dilihat bahwa titik-titik data House Prediction cenderung tidak menyebar mengikuti garis.

B. Association Rules

1. Analisis data

```
library(readxl)
datafix <- read_excel("Data Mining/datafix.xlsx")
View(datafix)
install.packages("arules")
install.packages("arulesViz")
library(arules)
library(arulesViz)
```

```

qq=as.matrix(datafix)
qq=as(qq,"transactions")
rules=apriori(qq,parameter=list(supp=0.2,conf=0.5,minlen=2))
rules=sort(rules, by="lift")
inspect(rules)
plot(rules,method="graph", control=list(type="itemsets"))

```

2. Interpretasi dari output

Untuk Association Rule Mining

```

> rules=apriori(qq,parameter=list(supp=0.2,conf=0.5,minlen=2))
Apriori

```

Parameter specification:

confidence	minval	smax	arem	aval	originalSupport
0.5	0.1	1	none	FALSE	TRUE

maxtime	support	minlen	maxlen	target	ext
5	0.2	2	10	rules	TRUE

Algorithmic control:

filter	tree	heap	memopt	load	sort	verbose
0.1	TRUE	TRUE	FALSE	TRUE	2	TRUE

Absolute minimum support count: 4

```

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[11 item(s), 20 transaction(s)] done [0.00s].
sorting and recoding items ... [9 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 done [0.00s].
writing ... [10 rule(s)] done [0.00s].
creating S4 object ... done [0.00s].

```

```

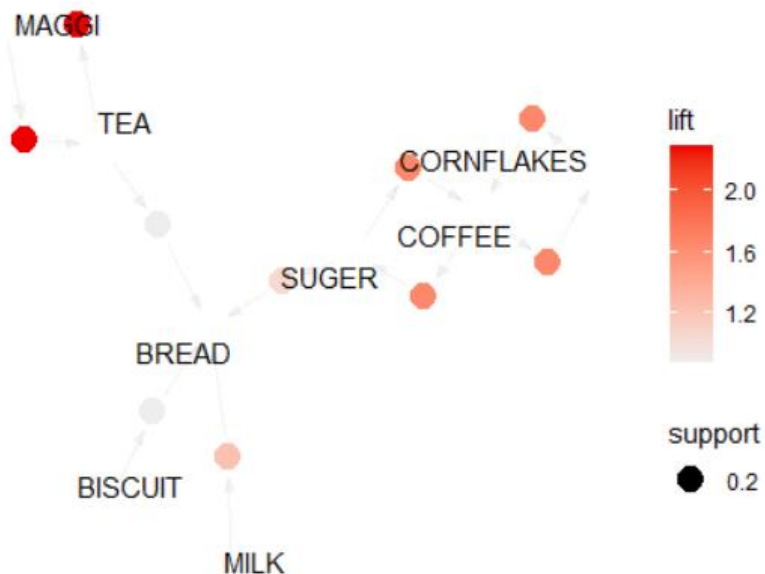
> rules=sort(rules, by="lift")
> inspect(rules)

```

	lhs	rhs	support	confidence
[1]	{MAGGI}	=> {TEA}	0.2	0.8000000
[2]	{TEA}	=> {MAGGI}	0.2	0.5714286
[3]	{COFFEE}	=> {SUGER}	0.2	0.5000000
[4]	{COFFEE}	=> {CORNFLAKES}	0.2	0.5000000
[5]	{SUGER}	=> {COFFEE}	0.2	0.6666667
[6]	{CORNFLAKES}	=> {COFFEE}	0.2	0.6666667
[7]	{MILK}	=> {BREAD}	0.2	0.8000000
[8]	{SUGER}	=> {BREAD}	0.2	0.6666667
[9]	{TEA}	=> {BREAD}	0.2	0.5714286
[10]	{BISCUIT}	=> {BREAD}	0.2	0.5714286

	coverage	lift	count
[1]	0.25	2.2857143	4
[2]	0.35	2.2857143	4
[3]	0.40	1.6666667	4
[4]	0.40	1.6666667	4
[5]	0.30	1.6666667	4
[6]	0.30	1.6666667	4
[7]	0.25	1.2307692	4
[8]	0.30	1.0256410	4
[9]	0.35	0.8791209	4
[10]	0.35	0.8791209	4

Untuk Visualisasi Association Rules



Interpretasi :

Jika dilihat dari scatter plot menunjukkan bahwa jika titik-titik tersebut lebih gelap itu artinya nilai peningkatan tinggi sedangkan yang titik lebih terang itu artinya nilai peningkatan rendah. Jadi, kita bisa menarik kesimpulan bahwa Maggi dan Tea memiliki peningkatan yang tinggi dari semua items.

BAB 5

KESIMPULAN

Dari decision tree diperoleh harga prediksi rumah yang paling mahal yaitu sebesar 1.425.678,938 dengan sebanyak 59 rumah. Kemudian harga prediksi rumah paling murah yaitu 307.010 dengan total rumah sebanyak 458. Sedangkan dari association tree diperoleh Maggi dan Tea memiliki peningkatan yang tinggi dari semua items.

DAFTAR PUSTAKA

- Arviana, Geofanni Nerissa. (2020). Coba Metode Decision Tree bagi Kamu yang Sulit Ambil Keputusan. <https://glints.com/id/lowongan/decision-tree-adalah/#.Yk9iEbf7Nkw>. Diakses pada 8 april 2022 pukul 09.00
- Haryati, Sudarsono, dan Suryana. (2015). Implementasi Data Mining Untuk Memprediksi Masa Strudi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Univversitas Dehasen Bengkulu), Jurnal Media Infotama Vol. 11 No. 2, September 2015, ISSN 1858 – 2680.
- Hermawan, A., Sukma, A. R., & Halfis, R. (2019). Analisis Algoritma Klasifikasi C 4.5 Untuk Memprediksi Keberhasilan Immunotherapy Pada Penyakit Kutil. Jurnal Teknik Komputer, 5(2), 155-160.
- Kusrini, E. T. L., & Taufiq, E. (2009). Algoritma data mining. Yogyakarta: Andi Offset.
- Mardi. (2016). Data Mining Klasifikasi Menggunakan Algoritma C4.5, Jurnal Edik Informatika, ISSN: 2407-0491.
- Pramadhani dan Setiadi. (2014). Penerapan Data Mining Untuk Klasifikasi Prediksi Penyakit ISPA (Infeksi Saluran Pernapasan Akut) Dengan Algoritma Decision Tree (ID3), Jurnal Sarjana Teknik Informatika Volume 2 Nomor 1, Februari 2014, eISSN: 2338-5197
- Tampubolon, K., Saragih, H., Reza, B., Epicentrum, K., & Asosiasi, A. (2013). Implementasi Data Mining Algoritma Apriori pada sistem persediaan alat-alat kesehatan. Majalah Ilmiah Informasi dan Teknologi Ilmiah, 1(1), 93-106.