

**ALGORITMA ASOSIASI DENGAN ALGORITMA APRIORI UNTUK
ANALISA ASOSIASI PENJUALAN**



**UNIVERSITAS
TEKNOLOGI
SUMBAWA**

DISUSUN OLEH :

VENI SEPTIANI	(20.01.013.016)
NOVITASARI	(20.01.013.012)
ISMED QALYUBI	(20.01.013.059)
ERNA PUTRI ATI	(20.01.013.024)
RIMA FAHRANA	(20.01.013.022)

**FAKULTAS REKAYASA SISTEM
PRODI TEKNIK INFORMATIKA
UNIVERSITAS TEKNOLOGI SUMBAWA**

2022

KATA PENGANTAR

Puji syukur kami ucapkan kehadiran Allah SWT yang telah melimpahkan rahmat, karunia, dan hidayah-Nya, sehingga kita sapat menyelesaikan makalah yang berjudul **“ALGORITMA ASOSIASI DENGAN ALGORITMA APRIORI UNTUK ANALISA ASOSIASI PENJUALAN PADA TOKO “SURYA USAHA”**

Penyusunan makalah ini digunakan untuk memenuhi tugas mata kuliah Kecerdasan Buatan yang dibimbing oleh Bapak Herfandi M.Kom. Pada kesempatan ini perkenankan kami mengucapkan terima kasih kepada: Bapak Herfandi M.Kom. selaku Dosen pengampu mata kuliah Kecerdasan Buatan. Pepatah mengatakan tidak ada gading yang tak retak. Oleh karena itu kami sadar dalam makalah ini masih terdapat kekurangan dan kesalahan, kami mohon maaf dan meminta kepada Bapak dosen, kiranya bersedia memberikan kritik dan saran untuk perbaikan selanjutnya. Sekian dari kami semoga tugas ini sesuai dengan apa yang diharapkan dan dapat bermanfaat bagi yang membacanya.

Sumbawa, 20 Desember 2021

Penyusun,

DAFTAR ISI

KATA PENGANTAR.....	ii
DAFTAR ISI	iii
DAFTAR TABEL	v
DAFTAR GAMBAR	vi
BAB I	1
PENDAHULUAN	1
1.1 LATAR BELAKANG	1
1.2 PERUMUSAN MASALAH.....	2
1.3 TUJUAN.....	2
1.4 MANFAAT.....	2
BAB II.....	3
LANDASAN TEORI.....	3
2.1 Data Mining.....	3
2.2 <i>Assosiation Rules</i>	7
2.3 Algoritma FP-Growth	8
2.4 Crisp-Dm	8
2.5 Dataset	10
2.6 RapidMiner	10
BAB III	11
METODOLOGI PENELITIAN	11
3.1 Metode Penelitian	11
3.2 Pengumpulan Data	11
3.2.1 Sumber Data	11
3.3 Pengelolaan Data Awal	15

3.3.1	Select Data.....	15
3.3.2	Cleaning.....	15
3.3.3	Transformation.....	16
3.3.4	Pemodelan	19
BAB IV	20
4.1	Pengujian dan Validasi Hasil	20
4.1.1	Proses Mining	20
	Hasil Proses Validation	28
BAB V	31
KESIMPULAN	31
5.1	Kesimpulan	31
5.2	Saran.....	31
DAFTAR PUSTAKA	33

DAFTAR TABEL

<i>Tabel 3. 1 dataset penjualan Toko “Surya Usaha”</i>	15
<i>Tabel 3. 2 tabel data clean</i>	16
<i>Tabel 3. 3 tabel data transformation</i>	18

DAFTAR GAMBAR

<i>Gambar 2. 1 rumus nilai suport</i>	8
<i>Gambar 2. 2 Rumus minimum confidence</i>	8
<i>Gambar 2. 3 Proses CRISP-DM.....</i>	9
<i>Gambar 4. 1 Tampilan Awal Rapidminer</i>	20
<i>Gambar 4. 2 tampilan menu proses RapidMiner</i>	21
<i>Gambar 4. 3 tampilan new proses pada RapidMiner</i>	21
<i>Gambar 4. 4 tampilan menu utama sebelum proses pada RapidMiner</i>	22
<i>Gambar 4. 5 tampilan proses read excel pada RapidMiner</i>	22
<i>Gambar 4. 6 tampilan import wizard pada RapidMiner</i>	23
<i>Gambar 4. 7 tampilan data yang akan diolah pada RapidMiner Kemudian pilihn next</i>	23
<i>Gambar 4. 8 Tampilan data yang akan diproses pada Rapiminer.....</i>	24
<i>Gambar 4. 9 Tampilan proses select attribut</i>	24
<i>Gambar 4. 10 Tampilan proses Numerical to Binominal</i>	25
<i>Gambar 4. 11 Tampilan Proses Remap Binominal</i>	26
<i>Gambar 4. 12 Tampilan proses FP-Growth</i>	27
<i>Gambar 4. 13 Tampilan proses Create Association Rules</i>	27
<i>Gambar 4. 14 Tampilan Connection proses</i>	28
<i>Gambar 4. 15 Tampilan hasil Frequentsets FP-Growth</i>	29
<i>Gambar 4. 16 Tampilan hasil Association Rules</i>	30

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Semakin banyaknya persaingan dalam dunia bisnis khususnya dalam industri penjualan menuntut pihak manajemen untuk putar otak mencari peluang untuk tetap berkecimpung dalam bisnisnya. Untuk menyiasati hal tersebut, tentunya pihak manajemen harus mampu menganalisa suatu data yang ada untuk di jadikan bahan acuan untuk menganalisa data tersebut.

Pemanfaatan data yang ada di dalam sistem informasi untuk menunjang kegiatan pengambilan keputusan, tidak cukup hanya mengandalkan data operasional saja, diperlukan suatu analisis data untuk menggali potensi- potensi informasi yang ada. Para pengambil keputusan berusaha untuk memanfaatkan gudang data yang sudah dimiliki untuk menggali informasi yang berguna membantu mengambil keputusan, hal ini mendorong munculnya cabang ilmu baru untuk mengatasi masalah penggalian informasi atau pola yang penting atau menarik dari data dalam jumlah besar, yang disebut dengan data mining.

Surya Usaha merupakan salah satu toko yang bergerak dalam bidang penjualan sembako yang berada di daerah Buer. Berbagai macam jenis kebutuhan sehari-hari khususnya sembako ditawarkan di toko tersebut dengan berbagai macam variasi harga tergantung dari jenis produk tersebut. Proses pencatatan transaksi yang dilakukan masih mengandalkan nota atau faktur untuk bukti transaksi, Data transaksi yang sudah dilakukan sampai pada proses pembuatan nota atau faktur kemudian di rekap oleh bagian administrasi untuk dilaporkan pada pemilik toko tersebut. Banyaknya data tersebut membutuhkan sebuah metode atau teknik yang dapat merubah banyaknya data agar dapat dimanfaatkan menjadi sebuah informasi berharga atau pengetahuan yang bermanfaat. Salah satu teknik yang dapat mengolah informasi dari kumpulan data yang besar menggunakan data mining.

Pada penelitian ini, penulis akan melakukan analisis terhadap data transaksi

yang sudah dilakukan dengan mekanisme perhitungan nilai *support* dan *confident* dari suatu hubungan item, yang nantinya hasil nilai *support* dan *confident* yang ada akan digunakan untuk proses pengambilan keputusan oleh pihak manajemen untuk meningkatkan strategi pemasaran.

1.2 PERUMUSAN MASALAH

Berdasarkan latar belakang yang telah dijelaskan diatas, masalah yang ditemukan adalah aturan asosiasi yang dihasilkan menjadi kurang akurat ketika jumlah data yang akan diolah besar, Untuk mengatasi permasalahan tersebut, dilakukanlah proses clustering data dengan menggunakan algoritma FP-Growth untuk menentukan rekomendasi produk pada dataset yang besar.

1.3 TUJUAN

Tujuan tugas akhir ini adalah menerapkan algoritma FP-Growth untuk clustering dalam menghasilkan rule rekomendasi produk pada jumlah dataset yang besar.

1.4 MANFAAT

1. Membantu menentukan kebijakan, khususnya dalam bidang strategi pemasaran.
2. Mempermudah pelanggan dalam berbelanja, karena rekomendasi produk yang diberikan lebih akurat.
3. Mencari perbandingan hasil sebelum dilakukannya cluster dan sesudah dilakukannya cluster dalam penerapan algoritma FP-Growth.

BAB II

LANDASAN TEORI

2.1 Data Mining

Data mining merupakan suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam basis data. Data mining juga merupakan proses yang menggunakan matematika, teknik statistik, kecerdasan buatan, dan machine learning untuk mengidentifikasi dan mengekstraksi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data besar.

Data mining merupakan bagian dari KDD yang terdiri dari beberapa tahapan, yaitu pemilihan data, pra-pengolahan, data mining serta interpretasi hasil. Sedangkan menurut Benni R Siburian (2014) dikutip dari Santoso (2007), data mining adalah proses yang bertujuan untuk menemukan suatu informasi atau pengetahuan yang berguna dari data yang berskala besar. Sehingga sering disebut sebagai bagian dari proses Knowledge Discovery in Database. Data mining menggunakan kekuatan dari komputer yang dikombinasikan dengan kemampuan bawaan manusia dalam menggambarkan visualisasi pola yang jelas. Dengan mengotomatisasi data mining, komputer menemukan pola dan trend data, ketika seseorang memanfaatkan penemuan- penemuan ini guna memutuskan pola yang benar-benar relevan (Benni R Siburian, 2014 dikutip dari Pramudiono, 2007).

Ada juga yang berpendapat Data Mining (DM) adalah inti dari proses knowledge discovery of database, melibatkan kesimpulan algoritma yang mencari data, mengembangkan model dan menemukan pola-pola yang sebelumnya tidak diketahui.

Siklus hidup proyek *data mining* menurut *Cross-Industry Standart Proses for Data Mining*(CRISP-DM) yang dikembangkan tahun 1996 terbagi dalam 6 fase (Kusrini, 2009). Berikut gambar dari Siklus hidup proyek *data mining*

1. Fase Pemahaman Bisnis (*Business Understanding Phase*)
 - a. Penentuan tujuan proyek dan kebutuhan secara detail dalam lingkup bisnis atau unit penelitian secara keseluruhan.

- b. Menerjemahkan tujuan dan batasan menjadi formula dari permasalahan *data mining*.
 - c. Menyiapkan strategi awal untuk mencapai tujuan.
2. Fase pemahaman data (*Data Understanding Phase*)
- a. Mengumpulkan data.
 - b. Menggunakan analisis penyelidikan data untuk mengenali lebih lanjut data dan pencarian pengetahuan awal.
 - c. Mengevaluasi kualitas data.
 - d. Jika diinginkan, pilih sebagian kecil group data yang mungkin mengandung pola dari permasalahan.
3. Fase pengolahan data (*Data Preparation Phase*)
- a. Siapkan data awal, kumpulkan data yang akan digunakan untuk keseluruhan fase berikutnya. Fase ini merupakan pekerjaan berat yang perlu dilaksanakan secara intensif.
 - b. Pilih kasus dan variabel yang ingin dianalisis dan yang sesuai analisis yang akan dilakukan.
 - c. Lakukan perubahan pada beberapa variabel jika dibutuhkan.
 - d. Siapkan data awal sehingga siap untuk perangkat pemodelan.
4. Fase Pemodelan (*Modelling Phase*)
- a. Pilih dan aplikasikan teknik pemodelan yang sesuai.
 - b. Perlu diperhatikan bahwa beberapa teknik mungkin untuk digunakan pada permasalahan *data mining* yang sama.
 - c. Jika diperlukan, proses dapat kembali ke fase pengolahan data untuk menjadikan data ke dalam bentuk yang sesuai dengan spesifikasi kebutuhan teknik *data mining* tertentu.
5. Fase Evaluasi
- a. Pengevaluasi satu atau lebih model yang digunakan dalam fase pemodelan untuk mendapatkan kualitas dan efektivitas sebelum disebarkan untuk

digunakan.

- b. Menetapkan apakah terdapat model yang memenuhi tujuan pada fase awal.
- c. Menentukan apakah terdapat permasalahan penting dari bisnis atau penelitian yang tidak tertangani dengan baik.
- d. Mengambil keputusan yang berkaitan dengan penggunaan hasil dari *datamining*.

6. Fase Penyebaran

- a. Menggunakan model yang dihasilkan. Terbentuknya model tidak menandakan telah terselesaikannya proyek.
- b. Contoh sederhana penyebaran: pembuatan laporan.
- c. Contoh kompleks penyebaran: penerapan proses *data mining* secara paralel pada department lain.

Menurut Larose dalam bukunya yang berjudul "*Discovering Knowledge in Data: An Introduction to Data Mining*", *datamining* dibagi menjadi beberapa kelompok berdasarkan tugas/pekerjaan yang dapat dilakukan (Larose, 2005), yaitu :

1. Deskripsi

Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan baris data (*record*) lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

4. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

5. Pengklasteran (*Clustering*)

Pengklasteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas obyek-obyek yang memiliki kemiripan. Klaster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan *record* dalam klaster yang lain. Berbeda dengan klasifikasi, pada pengklasteran tidak ada variabel target. Pengklasteran tidak melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target, akan tetapi, algoritma pengklasteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan *record* dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan *record* dalam kelompok lain akan bernilai minimal.

6. Asosiasi

Tugas asosiasi dalam *data mining* adalah untuk menemukan atribut yang muncul dalam satu waktu. Salah satu implementasi dari asosiasi adalah *market basket analysis* atau analisis keranjang belanja, sebagaimana yang akan dibahas dalam penelitian ini.

Dalam buku “Data Mining – Meramalkan Bisnis Perusahaan, Feri Sulianta & Dominikus Juju, 2010” menyatakan bahwa secara terperinci, ada empat tahap yang dilalui dalam *Data Mining* antara lain :

1. Tahap pertama:

Precise statement of the problem (mendefinisikan permasalahan yang ingin diketahui). Misalnya ingin mengetahui apakah seorang customer berpotensi memiliki kredit macet, atau mengidentifikasi seorang customer apakah akan pindah ke kompetitor bisnis kita, dan lain sebagainya. Setelah menemukan pertanyaan bisnis yang perlu dijawab oleh data mining, selanjutnya tentukan tipe tugas dasar untuk menjawab pertanyaan bisnis tersebut. Tugas dasar yang menjadi

dasar algoritma data mining adalah klasifikasi, regresi, segmentasi, asosiasi dan korelasi.

2. Tahap kedua:

Initial Exploration (Mempersiapkan data yang menjadi sumber untuk *data mining* termasuk data “*cleaning*” untuk dipelajari polanya). Setelah menentukan definisi masalah, langkah berikutnya adalah mencari data yang mendukung definisi masalah. Menentukan porsi data yang digunakan men-training data mining berdasarkan algoritma data mining yang telah dibuat. Setelah persiapan data selesai dilakukan, langkah berikutnya adalah memberikan sebagian data kedalam algoritma data mining.

3. Tahap Ketiga:

Model building and validation. Validasi apakah data mining memberikan prediksi yang akurat. Setelah *training* data selesai dilakukan, data mining tersebut perlu di-“uji” atau di-validasi keakuratannya terhadap data testing.

4. Tahap Ke-empat:

Deployment. Tahap ini memilih aplikasi yang tepat terhadap *data mining* untuk membuat prediksi

2.2 *Assosiation Rules*

Association rule merupakan suatu proses pada data mining untuk menentukan semua aturan asosiatif yang memenuhi syarat minimum untuk support (minsup) dan confidence (minconf) pada sebuah database. Kedua syarat tersebut akan digunakan untuk interesting association rules dengan dibandingkan dengan batasan yang telah ditentukan, yaitu minsup dan minconf.

Association Rule Mining adalah suatu prosedur untuk mencari hubungan antar item dalam suatu dataset. Dimulai dengan mencari frequent itemset, yaitu kombinasi yang paling sering terjadi dalam suatu itemset dan harus memenuhi minsup. Dalam tahap ini akan dilakukan pencarian kombinasi item yang memenuhi syarat minimum dari nilai support dalam database. Untuk mendapatkan nilai support dari suatu item A dapat diperoleh dengan rumus berikut:

$$\text{Support (A)} = \frac{\text{Jumlah transaksi yang mengandung Item A}}{\text{Total Transaksi}}$$

Kemudian, untuk mendapatkan nilai *support* dari dua item diperoleh dengan rumus berikut:

$$\text{Support (A, B)} = P(A \cap B) = \frac{\text{Jumlah Transaksi yang mengandung A dan B}}{\text{Total Transaksi}}$$

Gambar 2. 1 rumus nilai suport

Setelah semua frequent item dan large item set didapatkan, dapat dicari syarat minimumconfidence (mincof) dengan menggunakan rumus berikut :

$$\text{Confidence (A} \rightarrow \text{B)} = P(A|B) = \frac{\text{Jumlah Transaksi yang mengandung A dan B}}{\text{Jumlah Transaksi yang mengandung A}}$$

Gambar 2. 2 Rumus minimum confidence

2.3 Algoritma FP-Growth

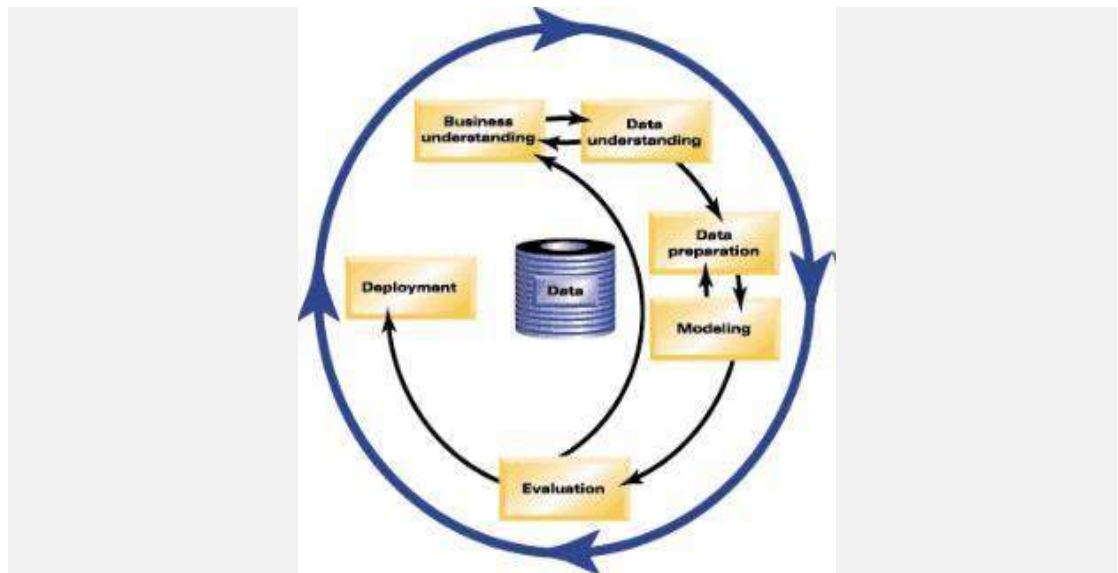
FP-Growth adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (frequent itemset) dalam sebuah kumpulan data. FP-growth menggunakan pendekatan yang berbeda dari paradigma yang selama ini sering digunakan, yaitu paradigma apriori.

2.4 Crisp-Dm

CRISP-DM memiliki kepanjangan Cross-Industry Standard Process for Data Mining adalah sebuah metode data mining yang dikembangkan bersama antara Daimler-Chrysler, SPSS, dan NCR dimana dari Namanya merupakan sebuah metode netral dan dapat digunakan dalam segala lini bisnis dan berbagai tool.

Sebagai sebuah metodologi, CRISP-DM menggambarkan fase dari tahapan — tahapan dalam sebuah proyek, pekerjaan yang terkait dalam tiap fase dan penjabaran terkait hubungan antar pekerjaan tersebut serta memberikan sebuah gambaran siklus hidup (life-cycle) dari Data Mining bila dilihat sebagai Model Proses.

Dari penggambaran tersebut metode ini memberikan sebuah proses standar yang bersifat umum atau tidak eksklusif dalam strategi pemecahan masalah dalam sebuah unit bisnis atau penelitian dengan menggunakan Data Mining yang sesuai atau tepat.



Gambar 2. 3 Proses CRISP-DM

Pada metode CRISP-DM ini memiliki 6 model tahapan seperti pada gambar 2 dalam keseluruhan proses data mining yaitu:

- a. Business/Research Understanding: Melakukan pengumpulan data perihal Business objective, penilaian terkait kondisi terkini, menetapkan tujuan dari proses data mining, dan mengembangkan rencana proyek.
- Data Understanding: Mengumpulkan data awal, deskripsi data, eksplorasi data, dan melakukan penilaian terkait kualitas data merupakan tahapan dalam fase ini. Dalam fase ini juga dilakukan eksplorasi data terkait ringkasan statistik yang dapat terjadi pada akhir fase ini serta melakukan clustering pada data untuk melihat pola data yang terbentuk.
- b. Data Preparation: Setelah data didapatkan perlu dilakukan proses sebuah proses seleksi, cleansing, dibuat dalam bentuk tertentu, dan di format sesuai kebutuhan.
- c. Modelling: Setelah data dibersihkan dan dibentuk sesuai kebutuhan kemudian

dibutuhkan sebuah modeling yang sesuai dan dikalibrasi perihal pengaturan agar didapatkan hasil optimal. Bila dibutuhkan kembali dapat dilakukan data preparation agar data dapat sesuai dengan teknik data mining yang dibutuhkan.

- d. **Evaluation:** Setelah didapatkan sebuah atau beberapa model sehingga dilakukan penilaian terkait kualitas dan efektifitas-nya. Kemudian ditentukan model seperti apa yang digunakan agar sesuai dengan objective pada fase 1 hingga diambil sebuah keputusan penggunaan dari hasil data mining.
- e. **Deployment:** Pada fase ini secara umum ada 2 aktifitas yang dilakukan yaitu Perencanaan dan monitoring hasil dari proses deployment serta melengkapi keseluruhan aktifitas sehingga menghasilkan laporan terakhir dan melakukan review dari proyek yang dilakukan.

2.5 Dataset

Pengertian dataset adalah sebuah kumpulan data yang berasal dari informasi-informasi pada masa lalu dan siap untuk dikelola menjadi sebuah informasi baru. Menurut Dayat Suryana dalam Controls Visual Basic Jilid 1, pengertian dataset adalah representasi data yang disimpan di memori dalam kondisi tidak terhubung (disconnected).

2.6 RapidMiner

RapidMiner adalah platform perangkat lunak ilmu data yang dikembangkan oleh perusahaan bernama sama dengan yang menyediakan lingkungan terintegrasi untuk persiapan data, pembelajaran mesin, pembelajaran dalam, penambahan teks, dan analisis prediktif.

BAB III

METODOLOGI PENELITIAN

3.1 Metode Penelitian

Pada penelitian ini, data yang digunakan adalah data penjualan barang pada Toko “Surya Usaha”. Data tersebut akan diolah menggunakan metode *algoritma FP-Growth*. Berikut tahapan dalam melakukan penelitian *data mining*.

1. Pengumpulan data

Tahapan ini menerangkan tentang darimana sumber data dalam penelitian ini didapatkan dan menemukan informasi yang bisa digunakan untuk penelitian.

2. Pengolahan data awal

Tahapan ini menerangkan tentang tahap awal dalam data mining. Pengolahan awal data meliputi proses input data keformat yang dibutuhkan, penggabungan data dan training data.

3. Metode yang diusulkan

Tahapan ini dijelaskan pemilihan dan penggunaan metode algoritma FP-Growth pada penelitian.

4. Pengujian dan Validasi hasil

Tahapan ini menjelaskan tentang pengujian, hasil pengujian akan di validasi dan kemudian di evaluasi. Penjelasan mengenai hal ini akan di paparkan pada BAB IV.

3.2 Pengumpulan Data

3.2.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data penjualan barang pada supermarket X dengan total 50 transaksi. Setaip baris menandakan bahwa data berada dalam satu transaksi yang sama

TID	Barang
1	Gula,Terigu,Telur
2	Teh,Rokok,Gula,Telur

3	Kopi,Keju,Mentos,Terigu
4	Teh,Mentos,Gula,Terigu
5	Kopi,Mentos,Gula,Terigu
6	Terigu,saos,minyak goreng,Kecap
7	Gula,saos,terigu,Kecap
8	Terigu,Kecap,saos,Minyak Goreng
9	terigu,gula,Kopi,Telur
10	Indomie,kecap,Telur,saos
11	Indomie,Teh Botol,Saos,Gula
12	Susu,Teh Botol,Indomie,Mentos
13	Teh Botol,Keju,Mentos,Roti
14	Gula,Teh,Susu,Aqua
15	Mentos,Teh botol, Indomie, Saos
16	Roti,Susu,Kopi,Telur
17	Detergen,Sabunmandi,Pasta gigi,Shampoo
18	Sabun Mandi,Aqua,Telur,Shampoo
19	Roti,Susu,Teh,Gula
20	Sabun mandi,Aqua,,Shampoo
21	Indomie,Aqua,Sabun Mandi, Pasta gigi
22	Minyak Goreng,Telur,pasta gigi
23	Kecap,Minyak Goreng,saos
24	minyak goreng ,Kecap,gula,terigu
25	Susu,teh,mentos,minyak goreng
26	Rokok,Korek,Teh,Gula
27	Rokok,Korek,Kopi,Gula
28	Aqua,Rokok,pasta gigi,Sabun Mandi
29	Susu,shampoo,saos,kecap,telur
30	Indomie,Roti,sisis
31	Teh,Mentos,Telur,Susu
32	Mentos,Teh botol, Indomie, Saos,Kecap

33	Mentos,Telor,sosis,saos,roti
34	aqua,sosis,saos,roti
35	Detergen,Pasta gigi,Sabun Mandi
36	Indomie, Saos,susu,telur
37	Gula,Indomie,Detergen,Aqua
38	Terigu,susu,sabun mandi,detergen
39	Aqua,gula,Kopi,teh,susu
40	susu,teh botol,terigu,Mentos
41	Detergen,pasta gigi,minyak goreng, telur
42	sosis,shampoo,telor,gula
43	Terigu,susu,mentos,roti
44	shampoo,sabun mandi,mentos,rokok
45	Sosis,Saos,Kecap,Teh Botol
46	korek,Indomie,Roti,Keju
47	indomie,rokok,Roti,Keju
48	Indomie,Keju,Telor,Sosis
49	Indomie,saos,kecap,Keju,Sosis
50	saos,detergen,sabun mandi
51	Roti,telur,rokok,susu
52	detergen,korek,sabun mandi,pasta gigi
53	mentos,roti,telur,rokok
54	keju,sosis,teh botol,rokok,telur
55	gula,Roti,keju
56	susu,telor,sabun mandi,Roti
57	Teh Botol,kopi,Mentos,detergen
58	roti,sabun mandi,susu,Keju
59	rokok,Gula,Terigu,korek
60	Keju,Terigu
61	Korek,gula,teh,Roti,Sosis
62	minyak goreng,roti,rokok,Shampoo
63	Detergen,roti,shampoo,mentos

64	Rokok,Teh Botol,susu,Keju
65	pasta gigi,Teh Botol,roti,Sosis
66	susu,sabun mandi,Terigu,Minyak goreng
67	korek,minyak goreng, Terigu, Keju
68	rokok,mentos,Roti,telur
69	detergen,rokok,Teh Botol,roti
70	kecap,saos,rokok,Teh Botol,Mentos
71	sosis,rokok,korek,Roti
72	Roti,susu,saos,sabun mandi
73	minyak goreng ,pasta gigi,telur,terigu
74	minyak goreng,Detergen,terigu
75	Rokok,Mentos, Terigu, tea
76	Rokok,korek,gula,Susu,Roti
77	detergen,pasta gigi,Susu,Sosis
78	Aqua,terigu,sosis,saos,Roti
79	Aqua,Roti,Teh Botol,susu
80	Sosis,Saos,Kecap,Aqua
81	Terigu,detergen,Kecap
82	Terigu,Susu,Keju,Minyak Goreng
83	Terigu,Telor,Saos, Minyak Goreng
84	Roti,sosis,keju,kecap
85	saos,Minyak Goreng,sosis,keju, Terigu
86	Minyak Goreng,Sosis,Keju,Susu
87	Roti,Keju,Gula,Kopi
88	Susu,Keju,Gula,Tea
89	Teh,kopi,keju,gula
90	Teh,Kopi,Keju,Susu
91	Minyak Goreng,,sosis,saos,Roti
92	Gula, Terigu, Sosis, Keju
93	Gula,Keju,Roti,Tea
94	Korek,Minyak Goreng,Saos,telur

95	Korek,Minyak Goreng,Terigu,Gula
96	Gula, Kopi, Roti,Minyak Goreng
97	Keju,Roti,Teh,Terigu,Telor
98	Kopi,Gula,Roti
99	Minyak Goreng,Saos,Sosis,Kecap
100	Teh,Kopi,Gula,Minyak Goreng

Tabel 3. 1 dataset penjualan Toko “Surya Usaha”

3.3 Pengelolaan Data Awal

Pengelolaan awal pada penelitian ini akan mencakup semua kegiatan untuk mempersiapkan data sebelum masuk proses pemodelan. Dalam melakukan pengolahan data awal, akan dilakukan beberapa tahapan agar pada akhirnya akan didapatkan data yang bisa digunakan pada tahap berikutnya. Tahapan tersebut antara lain: *select data*, *cleaning*, dan *transformation*

3.3.1 Select Data

Select data adalah proses mengambil data yang diperlukan untuk analisis data. Dari data penjualan yang diperoleh nantinya akan diambil atribut-attribut yang dibutuhkan untuk dijadikan parameter. Dari data tersebut parameter yang dibutuhkan untuk melakukan data mining adalah semua itemset/barang yang diperjual-belikan dalam transaksi

3.3.2 Cleaning

Pada tahap ini akan dilakukan proses pembersihan data untuk memastikan data yang telah dipilih tersebut telah layak untuk dilakukan proses pemodelan. Tahapan ini antara lain memperbaiki data yang rusak, membersihkan dan menghapus data yang tidak diperlukan.

Pada tahap ini dilakukan proses menghilangkan data atribut yang sama sehingga nantinya akan mendapatkan atribut data yang bersifat unik atau tidak memiliki data duplikat.

Roti
Gula
Terigu
Susu
Telur
Saos
Keju
Minyak Goreng
Sosis
Mentos
Rokok
Teh
Kecap

Tabel 3. 2 tabel data clean

3.3.3 Transformation

Tahap transformasi data adalah tahap mengubah format data yang telah dipilih keformat yang dimengerti oleh proses data mining. Dari parameter yang telah diperoleh selanjutnya kita masukkan data yang berasal dari data awal. Untuk setiap itemsets yang dibeli oleh pelanggan pada data penjualan kita masukkan angka 1 dan untuk itemsets yang tidak dibeli oleh pelanggan diberikan angka 0

Roti	Gula	Terigu	Susu	Telur	Saos	Keju	Minyak Goreng	Sosis	Mentos	Rokok	Teh	Kecap
1	0	0	1	0	0	0	0	0	0	0	0	1
0	1	0	0	1	0	0	0	0	0	0	1	0
0	0	1	0	0	1	0	0	1	0	0	0	0
0	0	0	1	0	1	0	0	0	0	0	1	1
0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	1	0	1	0	0	0	0	0	0	0	0
0	0	1	0	0	0	1	0	0	0	0	0	1
0	0	0	0	0	1	0	0	0	0	1	0	0
1	0	0	0	1	1	1	0	1	0	0	0	0
0	0	0	0	0	1	0	0	1	0	0	0	0
0	0	1	0	0	0	1	1	0	0	0	0	0
0	0	0	0	1	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	1	0	1	0	1	0
0	0	0	0	1	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	1	0
0	0	0	0	1	0	0	0	0	1	0	0	0
0	0	0	0	0	0	1	0	0	0	1	0	0
0	0	0	0	0	1	0	0	0	0	0	1	0
0	0	0	1	1	0	1	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0	0	1
0	0	1	0	0	1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	1	0	0	1	0
0	0	1	0	0	0	0	0	1	0	0	0	0
1	0	0	0	1	1	1	1	0	0	0	0	1
0	1	0	1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	1	0	0	0	0	0
1	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0	0	1
0	0	0	0	0	0	0	1	0	0	0	1	1
0	0	1	0	0	0	0	0	0	0	0	0	1

0	0	0	0	0	0	1	1	0	0	0	0	0
1	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	1	0
0	0	0	0	0	1	0	0	0	0	0	0	1
0	0	0	0	1	0	0	0	0	1	0	0	0
0	1	0	1	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	1	0	0
0	0	0	1	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	1	0	1	0	0
0	1	0	0	0	0	0	0	0	0	1	0	0
0	0	0	1	0	0	0	0	0	0	1	0	0
0	0	0	1	0	0	0	1	0	0	0	0	0
0	1	0	0	1	0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0	0	1	0	0	0
0	0	0	1	0	0	0	0	0	0	1	0	1
0	0	0	0	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	1	0	0
0	0	0	1	0	0	0	0	0	0	0	1	0
0	0	0	0	0	1	0	0	0	0	1	1	1
0	0	0	1	1	0	0	0	0	0	0	1	0

Tabel 3. 3 tabel data transformation

3.3.4 Pemodelan

Tahap ini peneliti menggunakan tools Rapidminer 9.10.001 untuk menganalisis data dengan pendekatan Asosiasi menggunakan algoritma FP-Growth. Pemilihan tools Rapidminer dianggap mampu digunakan untuk penelitian, prototyping, dan mendukung semua langkah proses data mining seperti persiapan data, proses pengolahan data, pembuatan rule asosiasi data, dll yang nantinya pembahasan mengenai ini akan dibahas di BAB IV

BAB IV

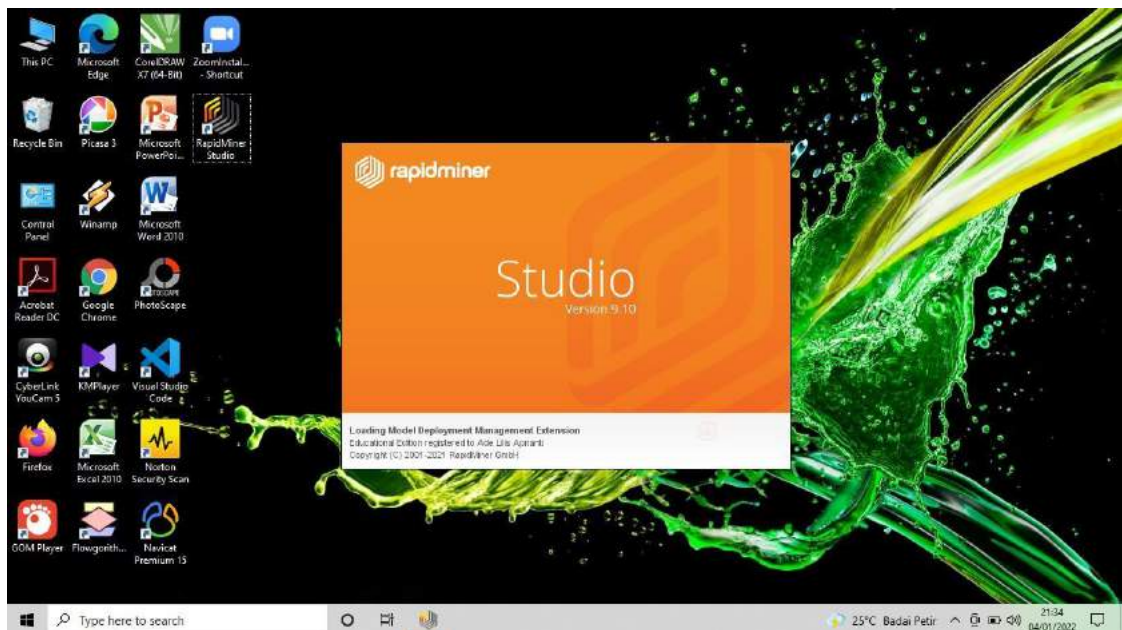
HASIL DAN PEMBAHASAN

4.1 Pengujian dan Validasi Hasil

4.1.1 Proses Mining

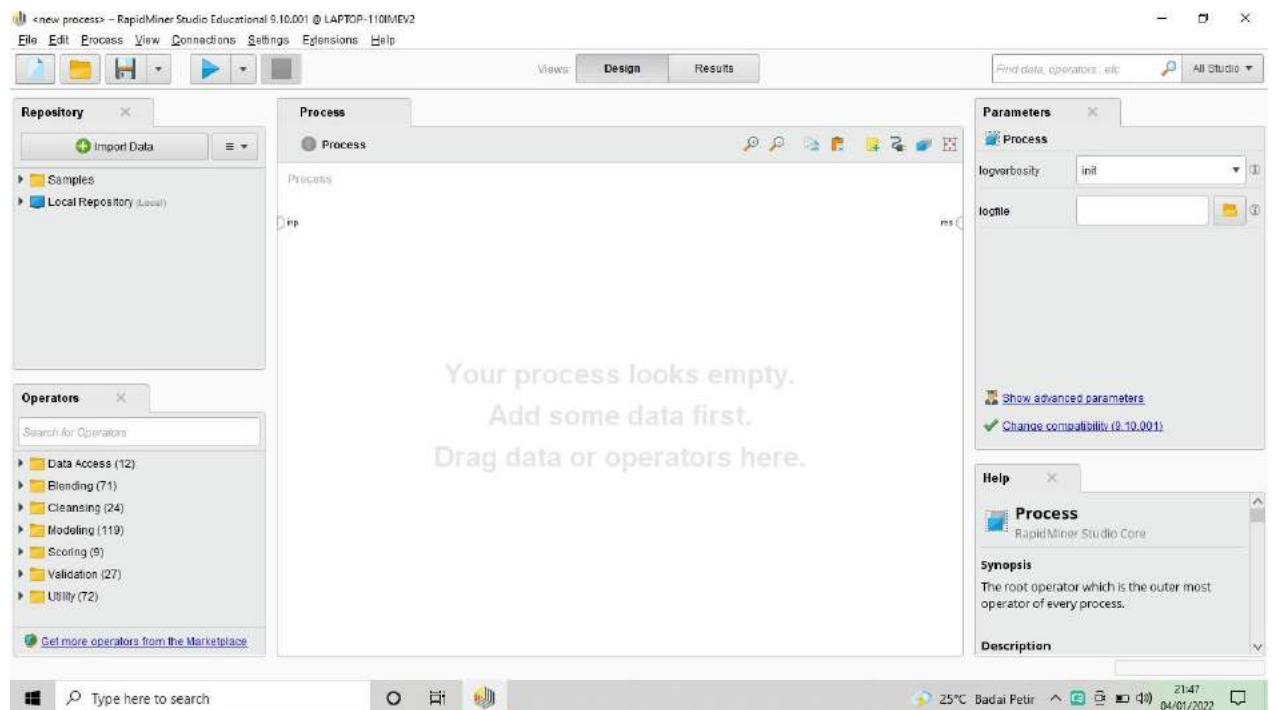
Pada tahap ini metode data mining untuk menemukan pengetahuan tersembunyi dan berharga dari data. Metode yang digunakan adalah assosiasi dengan algoritma FP-Growth. Berikut penerapan algoritma FP-Growth Memakai tool RapidMiner 9.10.001

1. Langkah pertama adalah membuka aplikasi *RapidMiner* akan muncul tampilan loading tampilan awal seperti berikut:



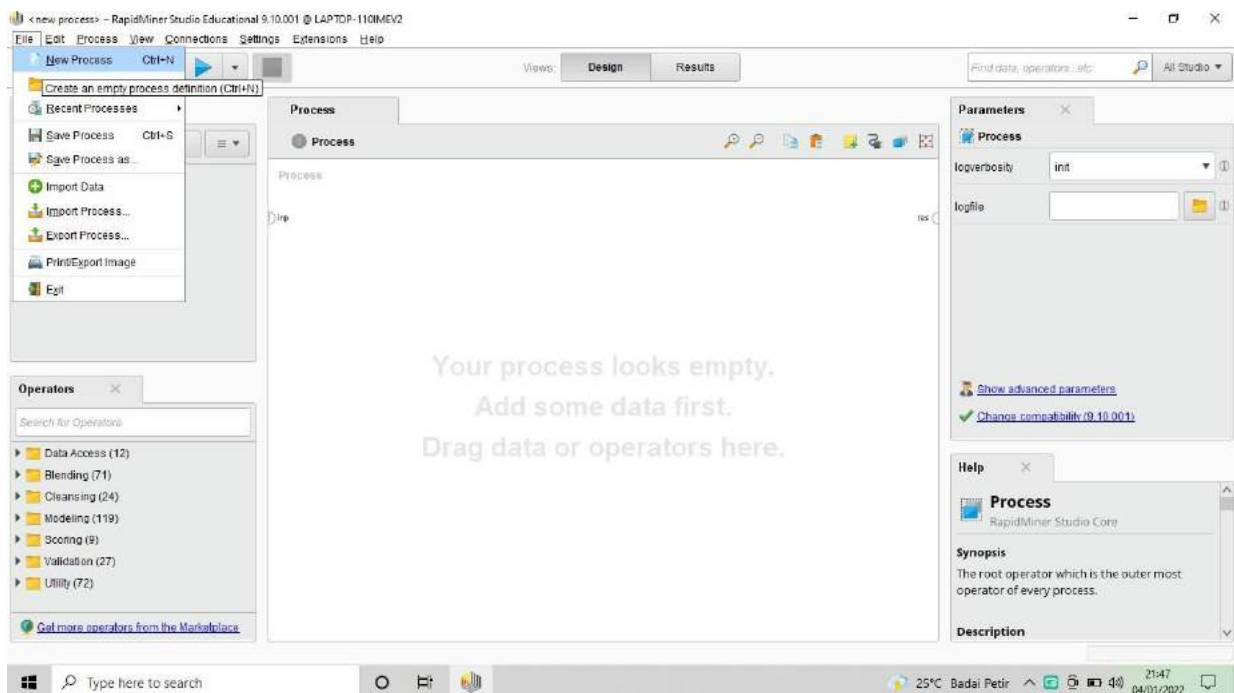
Gambar 4.1 Tampilan Awal Rapidminer

Setelah loading selesai akan muncul tampilan utama



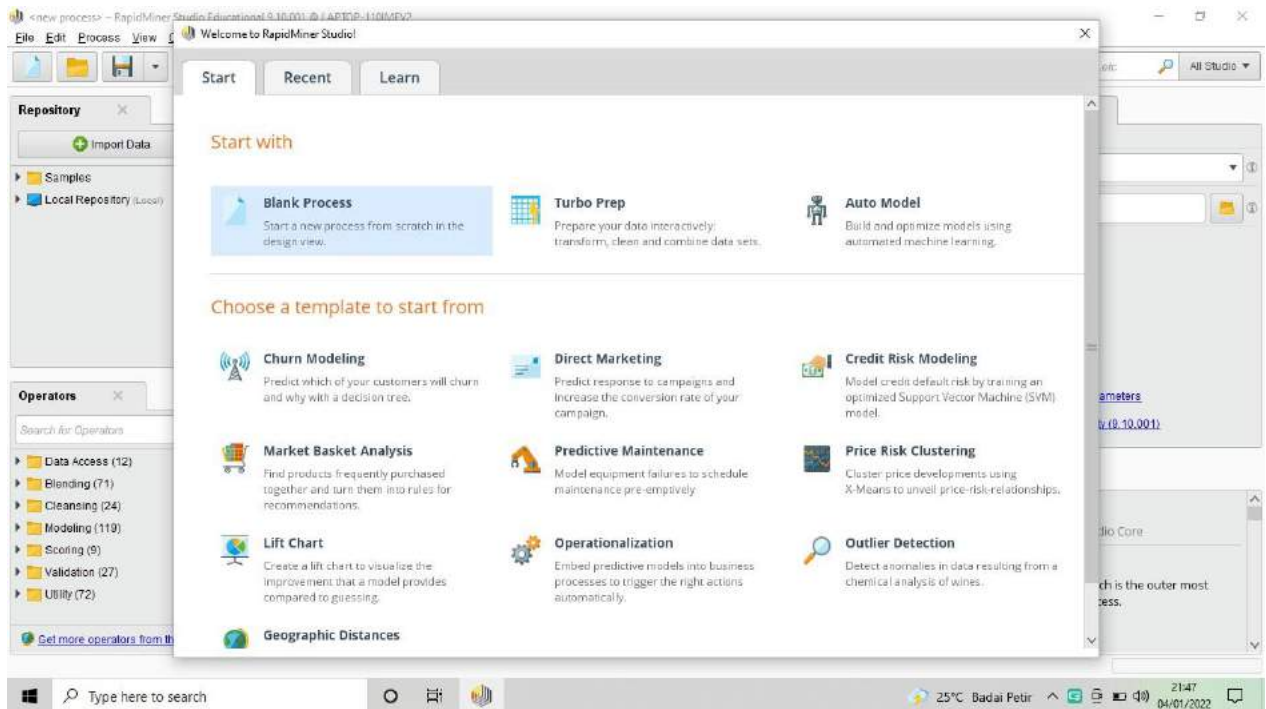
Gambar 4. 2 tampilan menu proses RapidMiner

Setelah muncul tampilan menu utama, tekan *New Process* untuk memulai proses pengolahan data.



Gambar 4. 3 tampilan new proses pada RapidMiner

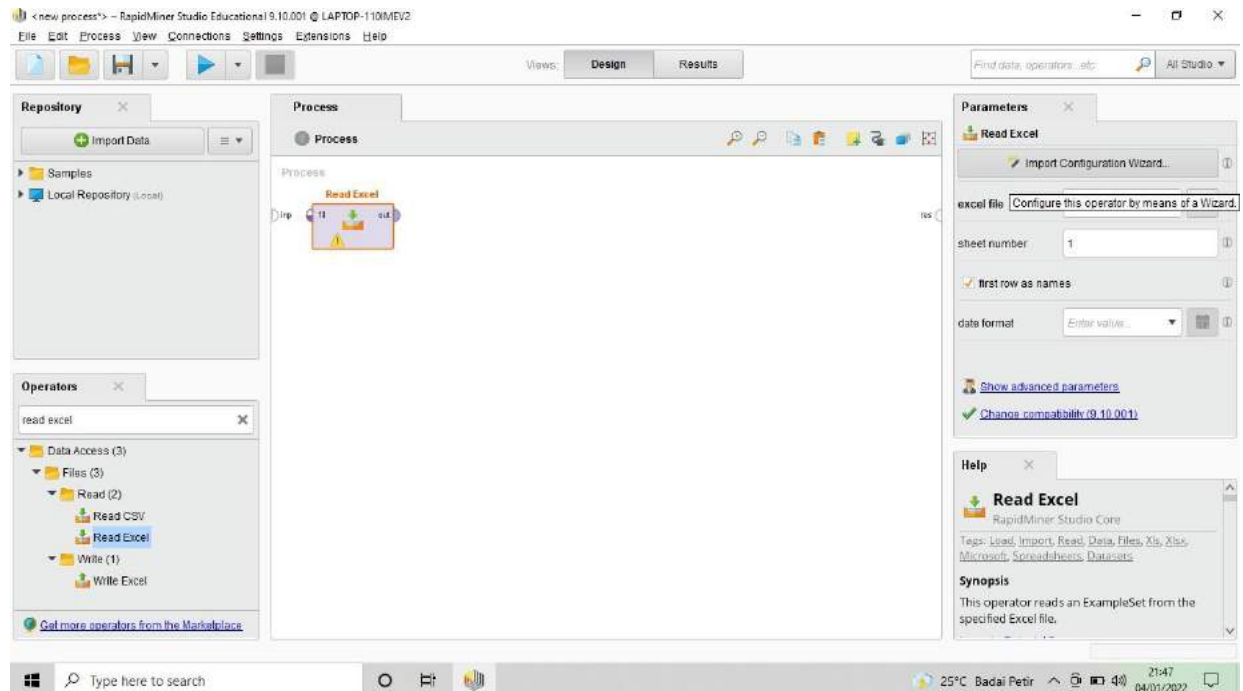
Kemudian pilih *Blank Proses*



Gambar 4. 4 tampilan menu utama sebelum proses pada RapidMiner

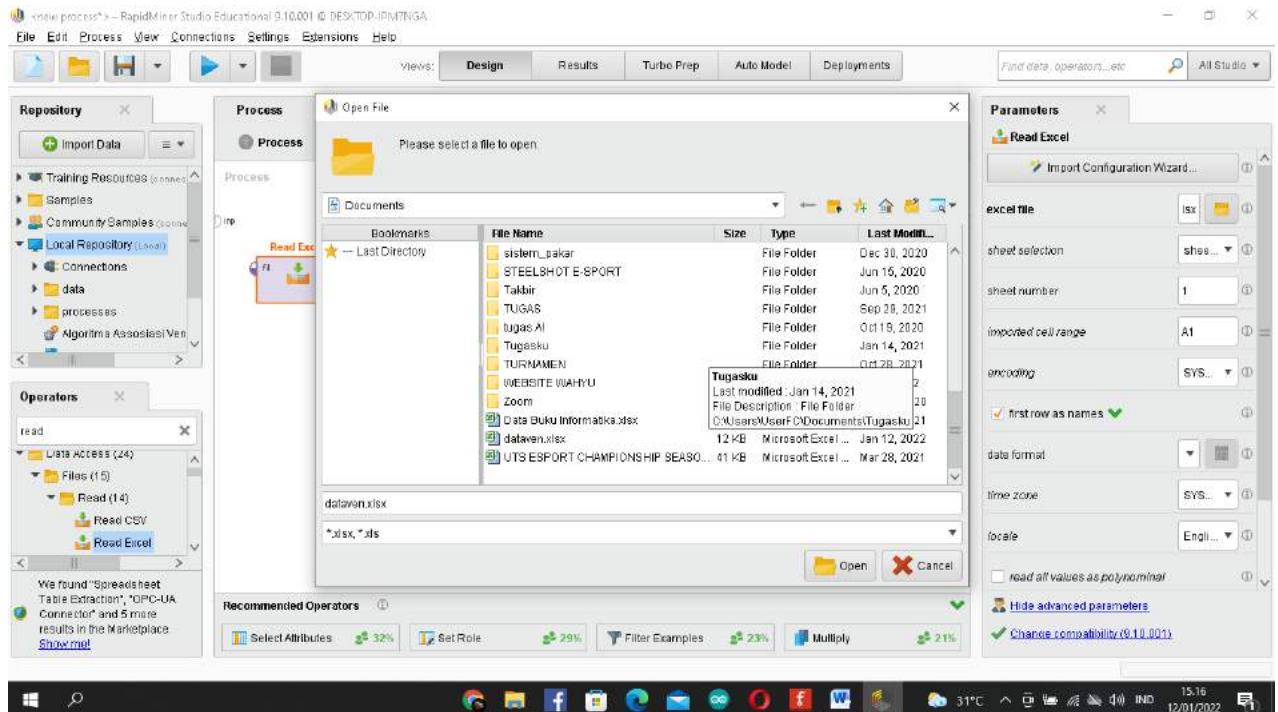
Setelah muncul tampilan process selanjutnya pada menu operator cari *Read Excel*

Dan pilih hingga muncul kotak *Read Excel* pada process



Gambar 4. 5 tampilan proses read excel pada RapidMiner

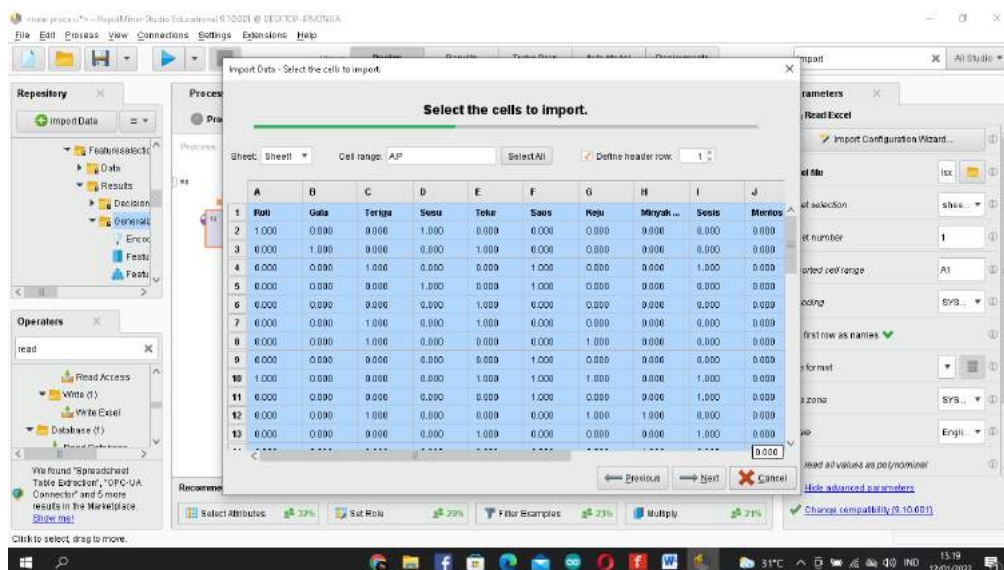
Kemudian klik *Import Configuration Wizard* untuk mengambil data uji transaksi dengan format data excel yang terdapat pada computer.



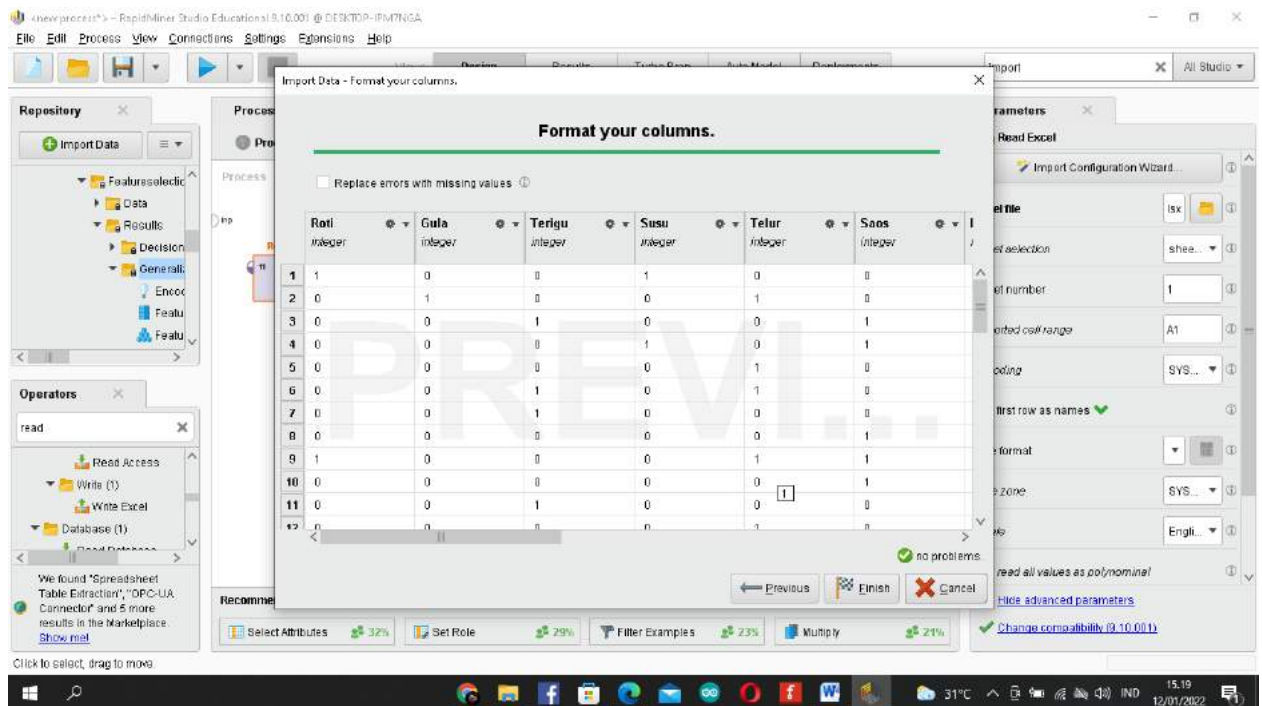
Gambar 4. 6 tampilan import wizard pada RapidMiner

2. Mengolah data

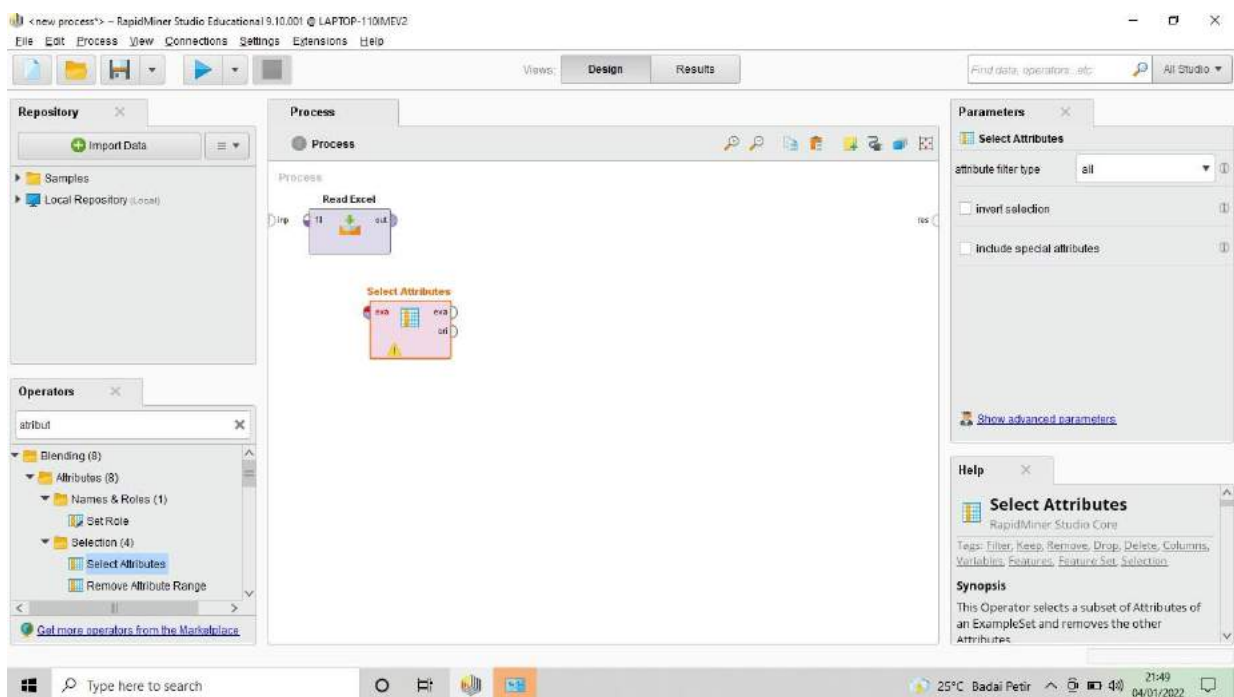
Menyiapkan data dari data yang telah siap untuk diolah. Pada data import wizard klik data yang telah diolah



Gambar 4. 7 tampilan data yang akan diolah pada RapidMiner Kemudian pilih next



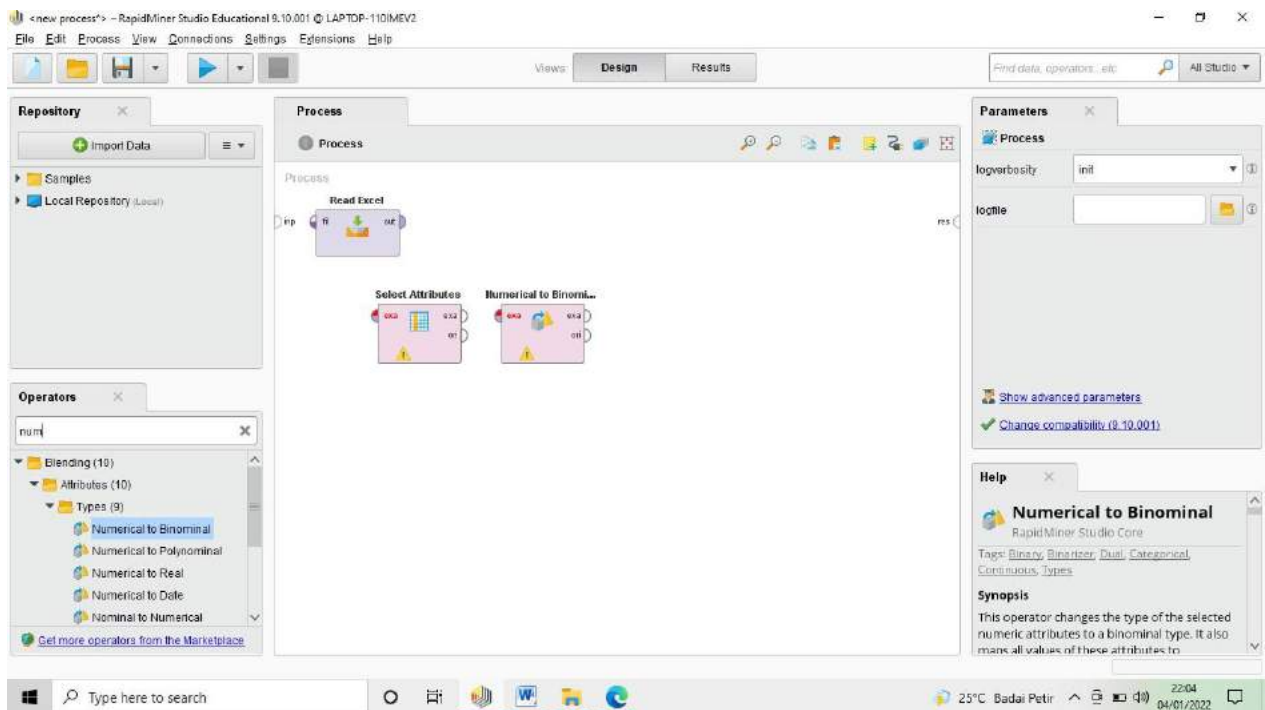
Gambar 4. 8 Tampilan data yang akan diproses pada Rapiminer



Gambar 4. 9 Tampilan proses select atribut

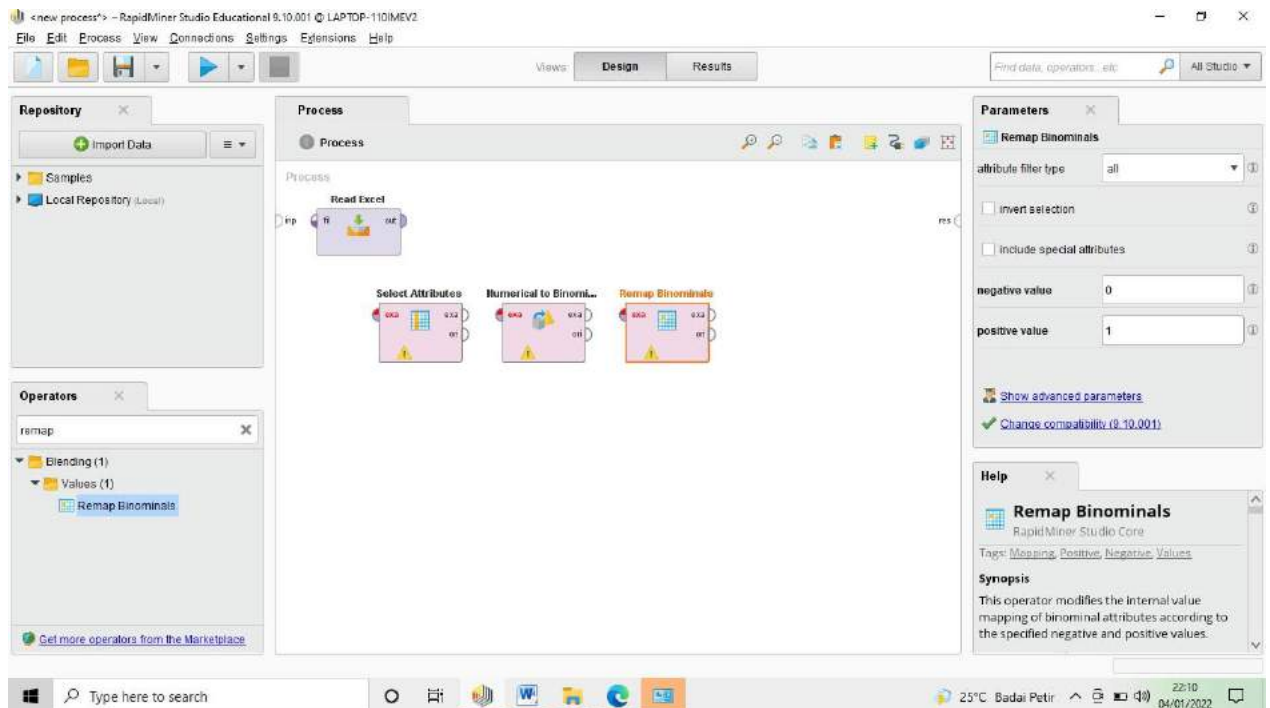
Kemudian pilih Finish. Setelah itu pada menu operator cari menu *Select Attributes*. Tahan dan seret menu *Select Attributes* ke bagian proses

Kemudian cari menu Numerical to binominals, tahan dan seret ke bagian proses



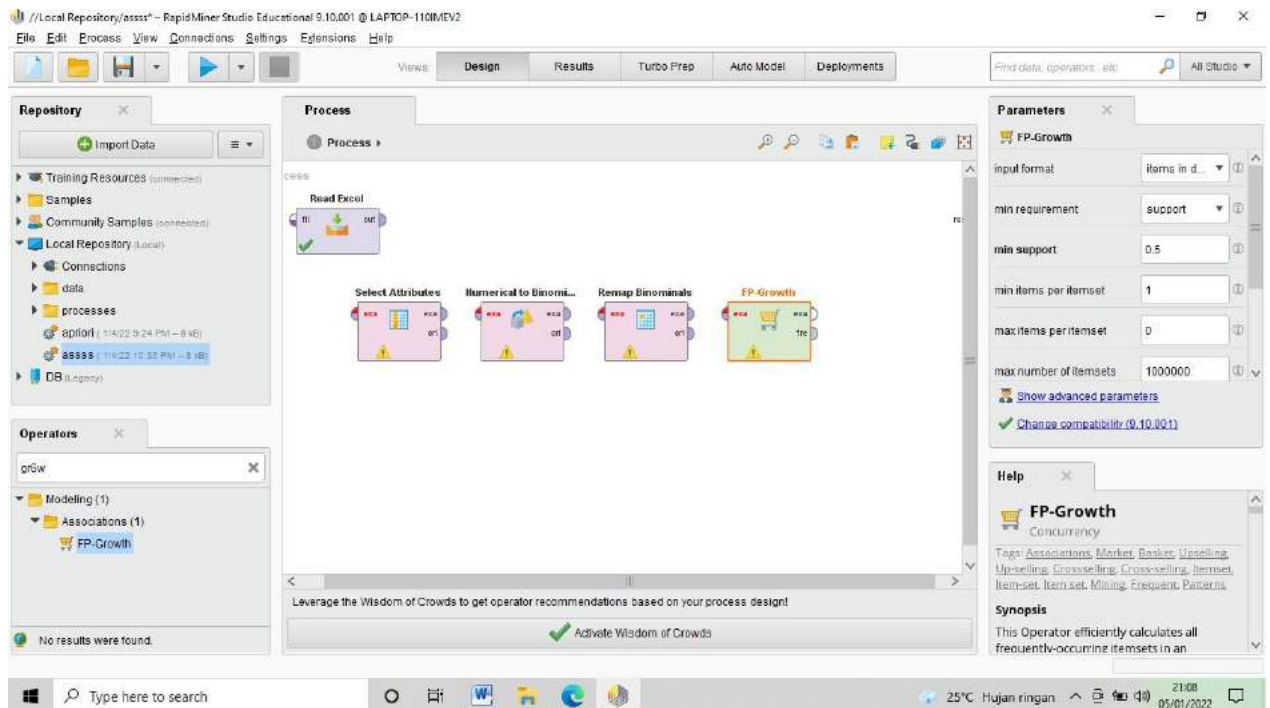
Gambar 4. 10 Tampilan proses Numerical to Binominal

Kemudian cari menu Remap binominals, tahan dan seret ke bagian proses.
Kemudian pada menu parameter tentukan negative value dan positive valuenya.
Masukkan angka 0 untuk negative valuenya dan angka 1 untuk positive valuenya



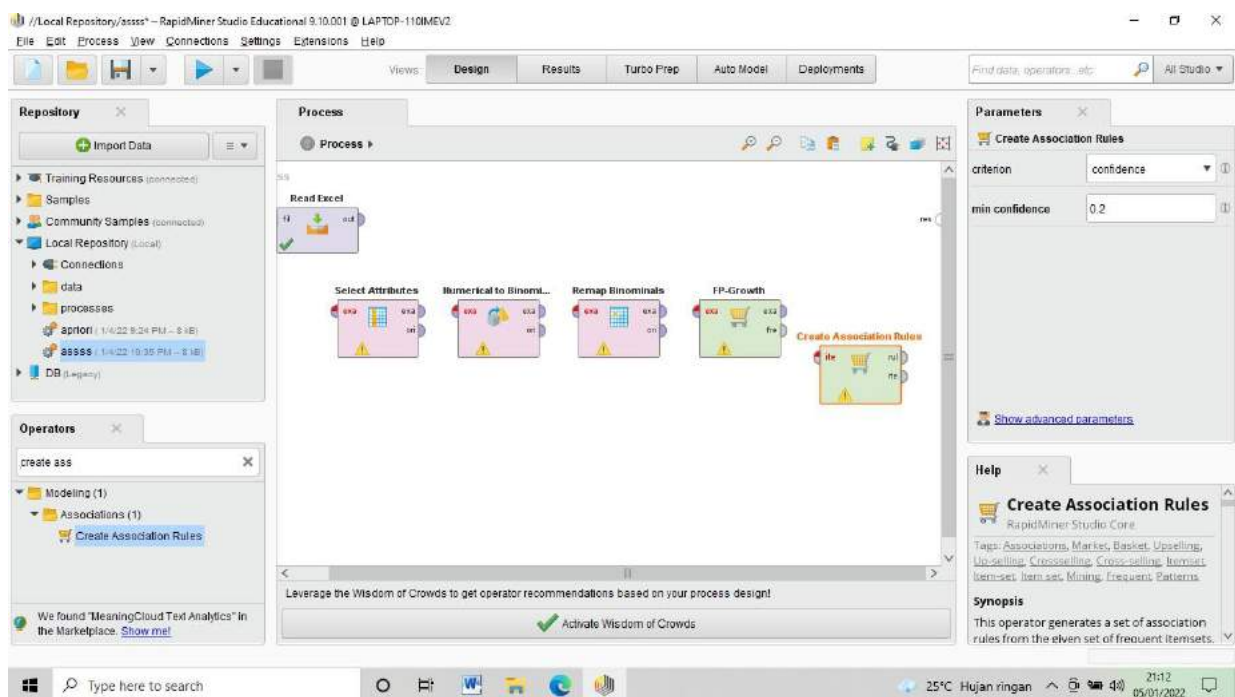
Gambar 4. 11 Tampilan Proses Remap Binominal

Kemudian cari menu FP-Growth, tahan dan seret ke bagian proses. pada menu parameter tentukan minimum requirement, min support, min items per itemsets, dan max items per itemsetnya. Untuk minimum requirement pilih support, dengan min supportnya kita masukkan 0.3, untuk min items per itemsetsnya 1, dan max items per itemsetnya 3



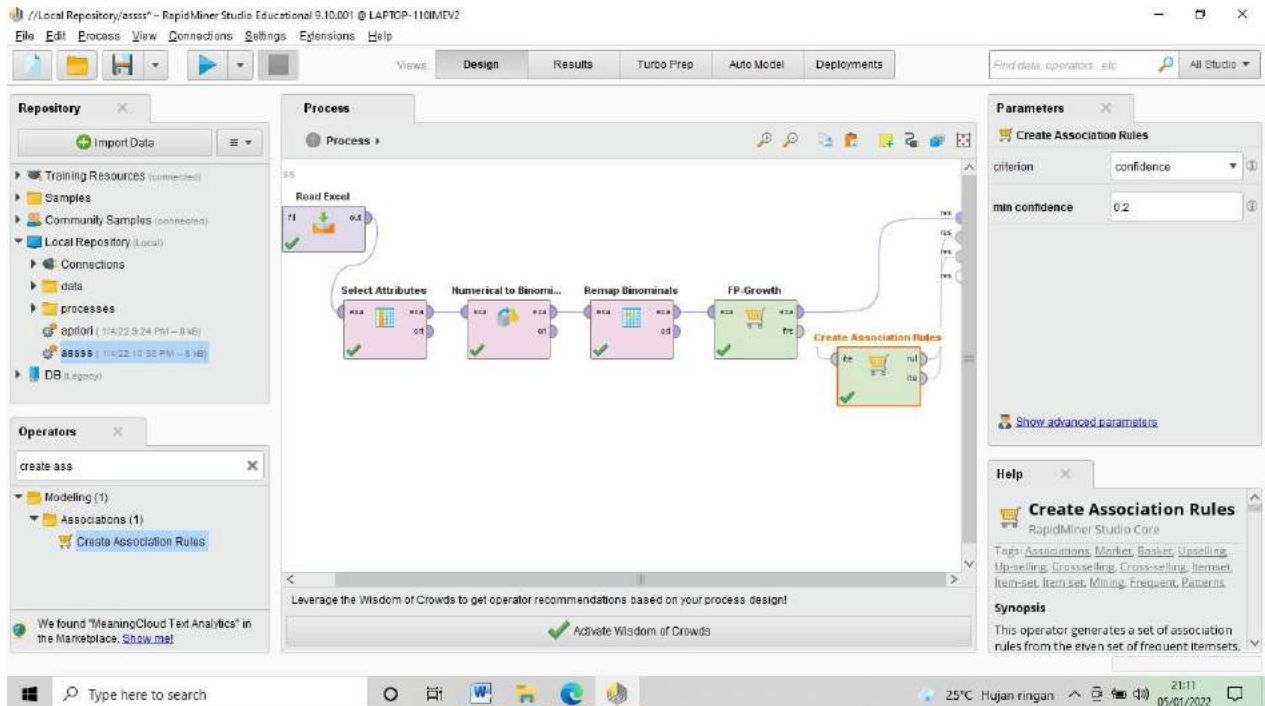
Gambar 4. 12 Tampilan proses FP-Growth

Kemudian cari menu Create association rules. Pilih confidence untuk criterionnya dan tentukan min confidennya 0.2



Gambar 4. 13 Tampilan proses Create Association Rules

Bentuk connection antar operator seperti pada gambar berikut



Gambar 4. 14 Tampilan Connection proses

kemudian klik run

Hasil Proses Validation

Hasil penentuan support dari masing-masing itemsets dengan rule yang telah ditentukan dengan algoritma FP-Growth yang terdapat pada data penjualan supermarket X.

Result History

FrequentItemSets (FP-Growth)

Size	Support	Item 1	Item 2
1	0.320	Telur	
1	0.260	Saos	
1	0.260	Susu	
1	0.240	Teh	
1	0.220	Kecap	
1	0.180	Rokok	
1	0.160	Keju	
1	0.160	Minyak Goreng	
1	0.140	Roti	
1	0.140	Sosis	
1	0.140	Terigu	
1	0.120	Mentos	
1	0.100	Gula	
2	0.080	Telur	Keju
2	0.080	Telur	Mentos

Gambar 4. 15 Tampilan hasil Frequentsets FP-Growth

Dari data tersebut dapat dilihat bahwa Telur memiliki nilai support 32%, Susu memiliki nilai support 26%, Teh memiliki nilai support 24%, Kecap memiliki nilai support 22%, Rokok memiliki nilai support 18%, Keju 16%, Minyak Goreng Memiliki nilai support 16%, Roti memiliki nilai support 14%, Sosis memiliki nilai support 14%, Sosis memiliki nilai support 14%, Terigu memiliki nilai support 14%, Mentos memiliki nilai support 12% dan untuk pasangan itemsets Telur dan keju memiliki nilai support 8%, pasangan itemsets Telur dan Mentos memiliki nilai support 8%, pasangan itemsets saos dan kecap memiliki support 12%, pasangan itemsets susu dan teh memiliki support 8%.

No.	Premises	Conclusion	Support	Confidence	Laplace	Gain	p-s	Lift	Conviction
3	Susu	Teh	0.080	0.308	0.867	-0.440	0.018	1.282	1.098
4	Teh	Susu	0.080	0.333	0.971	-0.400	0.016	1.282	1.110
5	Saos	Kecap	0.120	0.462	0.989	-0.400	0.063	2.098	1.449
6	Keju	Telur	0.080	0.500	0.931	-0.240	0.026	1.562	1.360
7	Kecap	Saos	0.120	0.545	0.918	-0.320	0.063	2.098	1.828
8	Mentos	Telur	0.080	0.667	0.964	-0.160	0.042	2.083	2.040

Gambar 4. 16 Tampilan hasil Association Rules

Hasil penentuan confidence dari data dengan rule asosiasi yang telah dibuat dari data menunjukkan:

- Jika membeli susu maka kemungkinan customer akan membeli teh adalah sebesar 30,8%
- Jika membeli teh maka kemungkinan customer akan membeli susu adalah 33,3%
- Jika membeli saos maka kemungkinan customer akan membeli kecap adalah 46,2%
- Jika membeli keju maka kemungkinan customer akan membeli telur adalah 50%
- Jika membeli kecap maka kemungkinan customer akan membeli saos adalah 54,5%
- Jika membeli mentos maka kemungkinan customer akan membeli telur adalah 66,7%

BAB V

KESIMPULAN

5.1 Kesimpulan

Dari uraian data pada bab-bab yang sudah dibahas sebelumnya dapat ditarik kesimpulan:

- Dari analisis dataset menggunakan algoritma FP-Growth diperoleh Telur memiliki nilai support 32%, Susu memiliki nilai support 26%, Teh memiliki nilai support 24%, Kecap memiliki nilai support 22%, Rokok memiliki nilai support 18%, Keju 16%, Minyak Goreng Memiliki nilai support 16%, Roti memiliki nilai support 14%, Sosis memiliki nilai support 14%, Sosis memiliki nilai support 14%, Terigu memiliki nilai support 14%, Mentos memiliki nilai support 12% dan untuk pasangan itemsets Telur dan keju memiliki nilai support 8%, pasangan itemsets Telur dan Mentos memiliki nilai support 8%, pasangan itemsets saos dan kecap memiliki support 12%, pasangan itemsets susu dan teh memiliki support 8%.
- Dari hasil perhitungan dataset yang telah diolah terdapat 6 pasangan itemsets yang tercipta dari rule association yang telah dibuat. Dari dataset tersebut didapatkan Jika membeli susu maka kemungkinan customer akan membeli teh adalah sebesar 30,8%, Jika membeli teh maka kemungkinan customer akan membeli susu adalah 33,3%, Jika membeli saos maka kemungkinan customer akan membeli kecap adalah 46,2%, Jika membeli keju maka kemungkinan customer akan membeli telur adalah 50%, Jika membeli kecap maka kemungkinan customer akan membeli saos adalah 54,5%, Jika membeli mentos maka kemungkinan customer akan membeli telur adalah 66,7%

5.2 Saran

Saran Agar penelitian ini terus berkembang, berikut saran-saran yang diusulkan :

1. Penelitian ini dapat dikembangkan dengan menggunakan metode aturan asosiasi rules data mining lainnya untuk dilakukan perbandingan.
2. Hasil dari penelitian ini dari rule yang terbentuk dapat dikembangkan

menjadi pengetahuan untuk sistem pendukung keputusan.

DAFTAR PUSTAKA

- Dawson, C. W. (2009). *Projects in Computing and Information System A Student's Guide*. England: Addison-Wesley.
- Gunadi - Penerapan Metode Data Mining Market Basket Analysis terhadap Penjualan Buku dengan Algoritma Apriori – 2012
- Han, J., & Kamber, M. , 2007, *Data Mining Concepts and Technique*. Morgan Kaufmann publisher
- Julsam dkk, Penambangan Data Format TextExcel dengan Software Tanagra, 2009.
- Kothari, C. R. (2004). *Research Methology methodes and Technique*. India: New Age Interntional.
- Kurniawati - Pemetaan Pola Hubungan Program Studi dengan Algoritma Apriori – 2014.
- Kusrini, dan Luthfi, E. T. (2007). *Algoritma Data mining*. Yogyakarta: Andi.
- Kusrini, dan Luthfi, E. T. (2009). *Algoritma Data mining*. Yogyakarta: Andi.
- Larose, D. T. (2005). *Discovering Knowledge in Data*. Canada: Wiley Interscience.
- Nurcahyo - Penerapan Data Mining dengan Algoritma Apriori untuk Mendukung Strategi Promosi Pendidikan
- Vercellis, C. (2009). *Business Intelligence : Data Mining and Optimization for Decision Making*. John Wiley dan Sons, Ltd
- Witten, H. I., Eibe, F., dan Hall, A. M. (2011). *Data Mining Machine Learning Tools and Techiques*. Burlington: Morgan Kaufmann Publisher.