

PROPOSAL PROYEK

12S4054– DATA MINING

**Market Basket Analysis Using FP-Growth
for Product Recommendation**



Disusun oleh:

Melani Basaria Pakpahan	12S17064
Juanda Antonius Pakpahan	12S17058
Yeni Chintya Panjaitan	12S17040

**PROGRAM STUDI SARJANA SISTEM INFORMASI
FAKULTAS INFORMATIKA DAN TEKNIK ELEKTRO
INSTITUT TEKNOLOGI DEL
NOVEMBER 2020**

Topik : *Market Basket Analysis*

Judul : *Market Basket Analysis for Product Recommendation*

Business Understanding

a. *Business Understanding*

Mengelola persediaan produk dalam sebuah bisnis adalah hal yang sangat dibutuhkan. Permintaan pelanggan yang berubah secara dinamis adalah masalah operasional *supermarket XYZ* yang sering terjadi. Perubahan tersebut sering membuat *supermarket* mengalami kesulitan dalam pengambilan keputusan terkait penyediaan stok suatu produk. Ketika *supermarket* menyiapkan pengadaan semua produk dengan jumlah yang sama, maka supermarket akan mengalami kerugian. Kerugian karena tidak sesuainya permintaan akan sebuah produk dengan persediaannya. Akibat lain yang juga ditimbulkan yaitu produk yang kurang diminati atau produk dengan frekuensi permintaan yang rendah tidak akan habis terjual dan jika terlalu lama akan memasuki masa kadaluarsa. Tingginya permintaan pelanggan terhadap persediaan yang sedikit akan menghasilkan kekecewaan dan tidak menutup kemungkinan pelanggan tidak kembali datang untuk membeli produk.

Salah satu cara yang dapat dilakukan untuk mengetahui kondisi pasar adalah dengan mengamati transaksi penjualan produk *supermarket*. Data transaksi penjualan yang disimpan dalam basis data dapat diolah lebih lanjut sehingga didapatkan informasi baru yang membantu perusahaan dalam mengambil keputusan terkait persediaan atau *stock* produk. *History* transaksi penjualan produk dapat digunakan untuk mengetahui frekuensi pembelian setiap produk.

Pada proyek ini, *market basket analysis* dilakukan dengan algoritma *fp-growth* untuk mengetahui hubungan atau pola-pola yang dihasilkan berdasarkan produk yang dibeli oleh pelanggan. *Market Basket Analysis* adalah solusi bagi *supermarket* untuk mengambil keputusan dalam menentukan strategi penjualan serta menyiapkan *stock* produk dengan tepat berdasarkan pola pembelian pelanggan. Hal tersebut dapat meningkatkan pelayanan, kepercayaan pelanggan terhadap *supermarket* dan meminimalisir kerugian. Dari hal tersebut *supermarket* mendapatkan informasi terkait produk apa saja yang memiliki

frekuensi pembelian yang tinggi maupun yang rendah, sehingga *supermarket* dapat dengan mudah mempersiapkan *stock* produk dengan efisien tanpa harus takut mengalami kerugian yang besar.

b. Situation Assesment

Supermarket XYZ merupakan toko yang menjual kebutuhan baik berupa makanan, minuman, dan lain sebagainya. Dalam pengoperasiannya toko ini masih menggunakan cara tradisional dalam penempatan tata letak produk dan penyediaan produk sehingga strategi pemasaran yang dilakukan oleh *supermarket XYZ* belum maksimal. Untuk meningkatkan strategi pemasaran yang tepat dapat digunakan *history* data penjualan pada *supermarket XYZ* untuk mengetahui hubungan atau pola- pola yang dihasilkan berdasarkan transaksi yang dilakukan oleh konsumen. Atribut *history data* yang dapat digunakan yaitu *order_id*, *user_id*, *order_number*, *order_dow*, *order_hour_of_day*, *days_since_prior_order*, *products*. Dengan mengetahui pola- pola yang dihasilkan dari transaksi produk yang terjual maka *supermarket XYZ* dapat mengetahui manakah produk yang sering dibeli oleh konsumen agar *supermarket XYZ* bisa lebih fokus pada produk tersebut baik dari segi tata letak dan penyediaanya.

c. Produce Project Plan

Pada pengerjaan proyek ini, algoritma *fp-growth* digunakan untuk menghasilkan sistem rekomendasi produk. Sebelum penggunaan algoritma tersebut, pertama yang perlu dilakukan yaitu menganalisis data transaksi penjualan *Supermarket XYZ*. Analisis data yang akan dilakukan yaitu pengecekan atribut dan tipe atribut yang terdapat pada dataset, selanjutnya akan dilakukan *data preparation* dengan menghapus atribut yang tidak dibutuhkan untuk pengerjaan proyek. Data yang telah dianalisis dan telah dipilih atribut yang diperlukan akan dilakukan proses *market basket analysis* dengan menggunakan algoritma *fp-growth* untuk menganalisa pola- pola yang dihasilkan dari transaksi. Algoritma *fp-growth* dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) pada kumpulan data. Selain itu, algoritma *fp-growth* menggunakan konsep struktur *tree* dalam pencarian *frequent itemset*. Karakteristik algoritma *fp-growth* merupakan struktur data menggunakan *tree* yang disebut dengan *fp-*

tree. Penggunaan *fp-tree* membuat algoritma *fp-growth* dapat langsung mengekstrak *frequent itemset* dari *fp-tree*.

Hasil dari proses *market basket analysis* menggunakan *fp-growth* diharapkan menghasilkan informasi dan pengetahuan berdasarkan rekomendasi yang dapat digunakan dalam pemenuhan stok produk dan tata letak produk. Selain itu, diharapkan juga dapat memberikan pengetahuan mengenai karakteristik dari *customer* berdasarkan transaksi yang telah dilakukan sebelumnya untuk membuat keputusan yang tepat.

Data Understanding

a. Data Understanding

- a. *Collect Initial Data*
- b. *Describe Data*
- c. *Explore Data*
- d. *Verify Data Quality*

Pada proyek ini, dataset yang digunakan yaitu, dataset yang berisi data transaksi pada sebuah *supermarket XYZ* yang diperoleh melalui *bigml.com*. Website tersebut merupakan portal penyedia dataset yang dikembangkan oleh BigML, Inc. Berikut adalah link datasetnya:

<https://bigml.com/dashboard/dataset/5fbdd95e2fb31c516b0031d1>

Terdiri dari 3,346,083 transaksi

Total row : 1.048.575

from 206,209 different users

33,819,106 products bought (49,685 different products)

Terdiri dari 7 *feature*:

- *order_id*: Order ID
- *user_id*: User ID
- *order_number*: Order number for a user set of orders
- *order_dow*: Order day of week (0 to 6)
- *order_hour_of_day*: Order hour of day (0 to 23)
- *days_since_prior_order*: Number of days since the previous order of the same user
- *products*: List of products bought in the order