

Penerapan Metode KNN-Regresi dan Multiplicative Decomposition untuk Prediksi Data Penjualan pada Supermarket X

Calvin Christopher Kurniawan, Silvia Rostianingsih, Leo Willyanto Santoso

Program Studi Informatika Fakultas Teknologi Industri Universitas Kristen Petra

Jl. Siwalankerto 121 – 131 Surabaya 60236

Telp. (031) – 2983455, Fax. (031) - 8417658

E-mail: calvinck00@gmail.com, silvia@petra.ac.id, leow@petra.ac.id

ABSTRAK

Supermarket X merupakan salah satu supermarket di Nusa Tenggara Barat yang memerlukan suatu cara untuk memprediksi penjualan di masa yang akan datang. Prediksi ini dibutuhkan oleh Supermarket X untuk memperkirakan rencana pembelian karena selama ini sering terjadi barang *stockout* ataupun *oversupply* yang menyebabkan kerugian bagi perusahaan.

Berdasarkan permasalahan yang terjadi, maka dalam penelitian ini dilakukan penerapan metode *KNN Regression* dan *Multiplicative Decomposition* dalam melakukan prediksi penjualan Supermarket X sehingga manajerial supermarket dapat merancang suatu strategi untuk melakukan penjualan di masa yang akan mendatang.

Hasil penelitian menunjukkan bahwa prediksi berdasarkan divisi, departemen, kategori, sub kategori, dan produk memiliki rata-rata tingkat *error* yang lebih kecil jika menggunakan metode *Multiplicative Decomposition* dengan $RMSE = 492,89$ dan $MAPE = 0,29$, sedangkan untuk metode *KNN Regresi* memiliki $RMSE = 757,77$ dan $MAPE = 0,36$.

Kata Kunci: *knn regresi, multiplicative decomposition, peramalan penjualan, forecasting, root mean squared error, mean absolute percentage error.*

ABSTRACT

Supermarket X is one of the supermarkets in West Nusa Tenggara that needs a way to predict sales in the future. This prediction is needed by Supermarket X to estimate the purchase plan because so far there have been frequent stockouts or oversupply which have caused losses to the company.

Based on the problems that occur, this study applies the *KNN Regression* and *Multiplicative Decomposition* methods in predicting Supermarket X sales so that supermarket managers can design a strategy to make sales in the future.

The results show that predictions based on divisions, departments, categories, sub categories, and products have a smaller average error rate when using the *Multiplicative Decomposition* method with $RMSE = 492.89$ and $MAPE = 0.29$, while the *KNN Regression* method has $RMSE = 757.77$ and $MAPE = 0.36$.

Keywords: *knn regression, multiplicative decomposition, forecasting, root mean squared error, mean absolute percentage error.*

1. PENDAHULUAN

Pada perkembangan zaman yang semakin berkembang ini, banyak perusahaan yang membutuhkan teknologi untuk mengolah data yang ada agar dapat mendukung jalannya bisnis. *Data Mining*

merupakan salah satu sistem yang berfungsi untuk mengubah data operasional, data transaksi, atau data lainnya pada perusahaan menjadi bentuk pengetahuan. *Data Mining* ini melakukan analisis data di masa lampau yang menghasilkan suatu pengetahuan untuk mendukung keputusan dan perencanaan organisasi [8]. Supermarket X merupakan salah satu supermarket di Nusa Tenggara Barat yang memerlukan suatu cara untuk memprediksi penjualan di masa yang akan datang. Prediksi ini dibutuhkan oleh Supermarket X untuk memperkirakan rencana pembelian karena selama ini sering terjadi barang *stockout* ataupun *oversupply* yang menyebabkan kerugian bagi perusahaan. Peramalan atau prediksi telah digunakan sebagai alat atau salah satu bahan pertimbangan dari pengambilan keputusan, terutama di sektor bisnis atau ekonomi di mana kerugian diminimalisir sebisa mungkin dan keuntungan atau laba diraih semaksimal mungkin [3]. Prediksi merupakan kegiatan mengidentifikasi suatu hal pada periode mendatang berdasarkan pada deskripsi orang lain/benda atau hubungan antara suatu hal yang dapat diketahui saat ini. Dalam melakukan prediksi terdapat dua teknik, yaitu dengan teknik kualitatif dan teknik kuantitatif [6].

KNN Regression merupakan algoritma KNN yang digunakan untuk melakukan regresi atau peramalan, metode ini telah dilakukan oleh beberapa penelitian sebelumnya. Metode *Multiplicative Decomposition* juga digunakan untuk dibandingkan tingkat akurasi dengan *KNN Regression*. Metode *Multiplicative Decomposition* akan memecah *dataset* menjadi beberapa musim atau *season* sehingga dapat dibaca bagaimana penjualan yang terjadi pada data yang bersifat *seasonal* dan memperhitungkan juga faktor tren dalam *dataset*. Metode *KNN Regression* dipilih karena memiliki kemampuan dalam memodelkan beragam masalah di berbagai bidang. *KNN* memiliki ketangguhan terhadap data yang banyak dan efektif apabila datanya besar seperti yang terdapat pada *dataset* di penelitian ini [11]. Algoritma *K-Nearest Neighbor* digunakan karena memiliki akurasi yang tinggi dengan rasio kesalahan kecil [7]. Metode *Multiplicative Decomposition* dipilih karena dapat meramalkan suatu data yang memiliki pola *seasonal*, tren, dan *cyclical*. Setiap *season* akan dihitung bobot penjualannya sehingga metode ini dapat meramalkan penjualan pada data yang bersifat *seasonal*. Terdapat juga *dashboard* yang akan diimplementasikan berbasis *website*. *Business Intelligence* merupakan serangkaian kegiatan untuk memahami situasi bisnis dengan melakukan berbagai jenis analisis pada data yang dimiliki oleh organisasi serta data eksternal pihak ketiga untuk membantu menentukan strategi, keputusan bisnis yang taktis, dan operasional dan mengambil yang diperlukan tindakan untuk meningkatkan kinerja bisnis [4].

2. PENELITIAN SEBELUMNYA

Penelitian akan menggunakan beberapa penelitian lain terkait yang telah dilakukan sebelumnya sebagai tinjauan studi. Berikut adalah penelitian-penelitian yang telah dilakukan sebelumnya:

2.1 Sistem Prediksi Pertumbuhan Jumlah Penduduk Kota Malang menggunakan Metode *K-Nearest Neighbour Regression*

Seruni, et al [13] melakukan penelitian untuk melakukan prediksi pertumbuhan jumlah penduduk kota Malang yang diharapkan dapat membantu pemerintah kota Malang dalam membuat rancangan perencanaan kota ke depannya. Penelitian ini menggunakan metode *KNN Regression*. Hasil penelitian ini adalah Nilai MAPE terendah yang berhasil didapatkan yaitu sebesar 0.02526% dengan fungsi jarak Euclidean, nilai $p = 6$, dan nilai $k = 5$ pada pengujian jangka pendek. Sedangkan pada pengujian jangka panjang untuk memprediksi jumlah penduduk beberapa bulan ke depannya, rata-rata nilai MAPE yang dihasilkan sebesar 0.13506% pada prediksi selama 24 bulan ke depan. Perbedaan penelitian ini dengan skripsi yang diusulkan adalah objek dari penelitian ini yaitu jumlah penduduk kota Malang. Penelitian ini juga menggunakan metode *KNN Regression* dengan metode pengujian yaitu, *root mean square error* (RMSE). Penelitian ini tidak menampilkan visualisasi dari hasil prediksi, sedangkan pada skripsi yang diusulkan menampilkan visualisasi dengan menggunakan *Dashboard Business intelligence* sehingga sistem mudah digunakan oleh pengguna.

2.2 LQ45 Stock Index Prediction Using *K-Nearest Neighbour Regression*

Tanuwijaya & Hansun [16] melakukan penelitian untuk memprediksi harga saham dengan benar. Penelitian ini menggunakan metode *KNN Regression*. Pengujian pada penelitian ini menghasilkan nilai *Mean Absolute Percentage Error* (MAPE) terbaik sebesar 0,02526%. Perbedaan penelitian ini dengan skripsi yang diusulkan adalah objek dari penelitian ini yaitu Indeks Saham LQ45. Selain itu penelitian ini tidak menampilkan visualisasi dari hasil prediksi, sedangkan pada skripsi yang diusulkan menggunakan metode *root mean square error* (RMSE) dan menampilkan visualisasi dengan menggunakan *Dashboard Business intelligence* sehingga sistem mudah digunakan oleh pengguna.

2.3 Penerapan Metode Regresi Linear untuk Prediksi Penjualan Properti pada PT.XYZ

Ayuni & Fitriana [1] melakukan penelitian untuk mendapatkan prediksi penjualan properti berdasarkan data penjualan properti pada PT XYZ. Regresi linear digunakan sebagai metode prediksi dengan kuantitas atau jumlah properti yang terjual sebagai variabel akibat dan periode penjualan properti sebagai variabel penyebab. Hasil yang diperoleh dari penelitian ini hanyalah berupa perkiraan berdasarkan data pada kurun waktu tertentu. Perbedaan penelitian ini dengan skripsi yang diusulkan adalah objek dari penelitian ini yaitu penjualan properti rumah. Penelitian ini juga menggunakan metode regresi linear dengan 3 metode pengujian yaitu, *mean square error* (MSE), *root mean square error* (RMSE), dan *mean absolute percentage error* (MAPE). Selain itu penelitian ini tidak menampilkan visualisasi dari hasil prediksi, sedangkan pada skripsi yang diusulkan menggunakan metode *root mean square error* (RMSE) dan menampilkan visualisasi dengan menggunakan

Dashboard Business Intelligence sehingga sistem mudah digunakan oleh pengguna.

2.4 Penerapan Metode *Multiplicative Decomposition* untuk Peramalan Barang pada PT. Agrinusa Jaya Sentosa

Kristiyanti & Sumarno [10] melakukan penelitian untuk memprediksi barang pada PT. Agrinusa Jaya Sentosa yang bergerak dibidang *animal health* dan *livestock equipment*. Dalam penelitian ini menggunakan data 1 tahun pada tahun 2018 dengan pendekatan metode *Multiplicative Decomposition*. Hasil penelitian diketahui bahwa hasil peramalan dengan metode *Multiplicative Decomposition (Seasonal) Average All* menghasilkan nilai MAD 1.124,87, MSE 1.995.484 dan MAPE 0,18 atau 18% dimana error MAPE tersebut tergolong kecil karena dibawah 20%. Perbedaan penelitian ini dengan skripsi yang diusulkan adalah jumlah metode yang digunakan hanya satu yaitu *Multiplicative Decomposition*, sedangkan penelitian yang akan dilakukan menggunakan 2 metode yaitu *Multiplicative Decomposition* dan *KNN Regression*. Dalam penelitian ini hanya dilakukan peramalan menggunakan *dataset* selama 1 tahun, sedangkan pada skripsi yang diusulkan akan menggunakan *dataset* lebih banyak yaitu selama 3 tahun agar perhitungan *seasonal* lebih akurat.

3. DATASET

Dataset yang digunakan dalam penelitian ini merupakan dataset yang berasal dari transaksi harian Supermarket X. *Dataset* yang digunakan bersumber dari *file excel* yang berisi data penjualan dari periode April 2019 – Februari 2022 yang kemudian akan diproses untuk menghasilkan prediksi dan *dashboard* informatif. Data akan dibagi menjadi data *training* dan data *testing*, dimana periode 2019 sampai 2021 akan dijadikan *data training* sedangkan periode Januari – Februari 2022 akan dijadikan sebagai *data testing*. *File excel* data penjualan kemudian akan di normalisasikan ke beberapa tabel yaitu *product*, *member*, *divisi*, *department*, *category*, *subcategory*, *transaksi*, dan *detail transaksi*.

4. METODE

Berikut adalah metode-metode yang digunakan dalam penelitian ini:

4.1 Regresi Linear

Analisis regresi adalah bersifat asimetri atau dua arah. Teknik regresi membuat prediksi nilai dengan nilai yang ada pada satu *variabel* (yang disebut *variabel* independen) pada *variabel* lain, yang disebut variabel dependen. Dalam hal ini tujuannya bukan bermaksud membuat prediksi yang sempurna. Dengan informasi pada independen bermaksud membuat prediksi nilai *variabel* dependen dengan *error* yang sekecil-kecilnya. Hubungan antar variabel tersebut dapat dimodelkan dalam sebuah persamaan. Proposisi yang digunakan pada analisis regresi yaitu dari variabel independen X dan variabel dependen Y dinyatakan dengan regresi Y pada X1 [14]. Regresi Linear merupakan solusi yang cocok untuk digunakan oleh perusahaan multiproduk karena dengan memperkirakan berbagai kombinasi produk, perusahaan dapat memaksimalkan keuntungan serta memperkirakan jumlah produksi yang tepat. Namun dalam menerapkan regresi linear diperlukan pengetahuan mengenai kombinasi produk yang tepat, target pasar, serta jumlah permintaan terhadap produk itu sendiri [9]. Rumus untuk Regresi Linear dengan metode kuadrat terkecil atau sederhana dapat dituliskan dalam bentuk sebuah persamaan sederhana adalah seperti persamaan (1),(2), dan (3).

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2} \quad (1)$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \quad (2)$$

$$y = a + b.x \quad (3)$$

dengan y adalah kuantiti penjualan, x adalah periode penjualan atau bulan penjualan, a adalah konstanta yang menunjukkan besarnya nilai y apabila $x = 0$, dan b adalah besaran perubahan nilai y .

4.2 Pembangunan Model KNN Regresi

K-Nearest Neighbor (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut [2]. KNN pada dasarnya digunakan untuk melakukan klasifikasi, tetapi dapat juga digunakan untuk melakukan regresi, yang biasa disebut juga sebagai *KNN Regression*, di mana hasil regresi didapatkan dari rata-rata k tetangga (*neighbor*) terdekat. Pada data *time series* dengan satu fitur, untuk melakukan perhitungan KNN, data tersebut perlu direpresentasi ulang terlebih dahulu ke dalam bentuk yang sesuai untuk algoritma *supervised learning* [5]. Diasumsikan sebuah *time series* $y = \{y_1, y_2, \dots, y_n\}$ dengan panjang n data, *time series* tersebut akan direpresentasikan dalam bentuk sebuah tabel dengan kolom fitur sejumlah p dan sebuah kolom label atau target. Untuk sebuah output atau keluaran $Target_t = y_t$, dengan $t > p$, maka fiturnya adalah $P_i = Y_{(t-p+i)}$, di mana $i = 1, 2, \dots, p$. Langkah-langkah algoritma *KNN Regression* untuk menghitung nilai prediksi adalah sebagai berikut [5]:

- Menentukan nilai p , k , dan fungsi jarak
- Representasikan ulang *time series* ke dalam bentuk model
- Menghitung jarak antara data yang akan diprediksi dengan data-data sebelumnya.
- Mengurutkan data-data sebelumnya berdasarkan jarak terkecil, kemudian menghitung rata-rata target dari k data teratas yang dapat dilihat pada persamaan (4). Hasil perhitungan rata-rata tersebut menjadi hasil prediksinya.

$$y(t+1) = \frac{1}{k} \sum_{i=1}^k y_i \quad (4)$$

4.3 Multiplicative Decomposition

Metode dekomposisi sering digunakan tidak hanya dalam menghasilkan ramalan, tetapi juga dalam menghasilkan informasi mengenai komponen deret berkala dan tampak dari berbagai faktor, seperti trend (*trend*), siklus (*cycle*), musiman (*seasonal*), dan keacakan (*irregular*) pada hasil Kristiyanti & Sumarno [10]. Menurut Syaifulloh [15], menyatakan bahwa, “Metode peramalan *multiplicative decomposition* juga terbagi menjadi 2 yaitu *Multiplicative Decomposition – Average All* dan *Multiplicative Decomposition – Centered Moving Average*”.

4.4 Pengukuran Keakuratan

Salah satu metode untuk pengukuran keakuratan adalah RMSE (*Root Mean Square Error*), RMSE merupakan alat seleksi model berdasarkan pada error hasil estimasi. Error yang ada menunjukkan seberapa besar perbedaan hasil estimasi dengan nilai yang akan diestimasi. Pengujian akurasi sistem dilakukan dengan menggunakan metode *Root Mean Square Error* (RMSE) untuk mencari nilai error berdasarkan perbandingan antara perhitungan manual dan hasil output pada sistem yang telah dibangun dimana semakin kecil (mendekati 0) nilai RMSE maka hasil prediksi akan semakin akurat [12]. Secara sistematis, RMSE dirumuskan seperti persamaan (5):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y - \hat{Y}_i)^2}{n}} \quad (5)$$

MAPE (*Mean Absolute Percentage Error*) adalah nilai rata-rata diferensial absolut antara nilai *forecast* dengan nilai aktual yang dinyatakan dalam angka prosentase sebagai nilai aktual. MAPE dihitung sebagai rata-rata diferensial absolut antara nilai yang diamalkan dengan nilai actual seperti persamaan (6).

$$\frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (6)$$

5. PENGUJIAN

Pada bab ini akan dijelaskan mengenai pengujian terhadap sistem yang sudah dibuat dan dikembangkan sebelumnya. Menu dan fitur yang terdapat pada halaman *website* akan dilakukan uji coba serta melakukan pengujian terhadap akurasi prediksi dengan metode *KNN Regresi* dan *Multiplicative Decomposition* menggunakan data tahun 2019-2021 untuk memprediksi tahun 2022 bulan Januari dan Februari. Tujuan dari uji coba yang dilakukan ini adalah untuk menentukan kombinasi nilai *input* parameter yang memberikan nilai RMSE dan MAPE terkecil pada setiap data uji coba. Dataset ujicoba yang digunakan adalah data transaksi untuk 8 departemen utama, 10 divisi dengan penjualan terbanyak, 10 *subcategory* dengan penjualan terbanyak, 10 *category* dengan penjualan terbanyak, dan 10 produk dengan penjualan terbanyak. Semua prediksi baik berdasarkan divisi, *department*, kategori, subkategori, maupun *product* memiliki cara yang sama hanya berbeda di dataset dan variabel saja. Oleh karena itu pada sub bab selanjutnya akan diberikan contoh untuk perhitungan prediksi dan akurasi berdasarkan *product* saja. Data dari tahun 2019 sampai dengan tahun 2021 akan diprediksi untuk menentukan penjualan di tahun 2022 bulan Januari dan Februari yang kemudian hasil dari prediksi tersebut akan dibandingkan dengan data penjualan yang asli.

5.1. Pengujian Nilai P dan K untuk KNN Regresi (Produk)

Metode *KNN-Regresi* membutuhkan 2 *input* utama yaitu nilai K (jumlah tetangga terdekat) yang kemudian akan dilakukan rata-rata dari *neighbour* terdekat tersebut. Nilai P adalah jumlah fitur yang akan menentukan suatu target data. Nilai dari K dan P akan dilakukan uji coba untuk menentukan nilai yang optimal. Uji coba akan dilakukan dengan menggunakan nilai K dan P sebanyak 1 sampai 3, contoh perhitungan akan menggunakan K senilai 2 dan P senilai 1. Tujuan dari uji coba yang dilakukan ini adalah untuk menentukan kombinasi nilai P dan K yang memberikan nilai RMSE dan MAPE terkecil pada setiap data uji coba. Tabel 1 merupakan daftar transaksi yang terjadi setiap bulannya pada *product* “INDOMIE GORENG/40”.

Tabel 1. Penjualan Produk Indomie Goreng/40 Per Bulan

No	Jumlah Penjualan	Periode
1	413	2019-April
2	279	2019-May
3	685	2019-June
4	303	2019-July
5	482	2019-August
6	313	2019-September
7	565	2019-October
8	593	2019-November
9	647	2019-December
10	747	2020-January

11	827	2020-February
12	875	2020-March
13	1162	2020-April
14	544	2020-May
15	346	2020-June
16	198	2020-July
17	486	2020-August
18	814	2020-September
19	334	2020-October
20	328	2020-November
21	512	2020-December
22	291	2021-January
23	307	2021-February
24	415	2021-March
25	457	2021-April
26	180	2021-May
27	511	2021-June
28	112	2021-July
29	235	2021-August
30	212	2021-September
31	309	2021-October
32	258	2021-November
33	351	2021-December
34	502	2022-January
35	436	2022-February

Prediksi bulan Januari 2022 akan melihat data 1 bulan sebelumnya karena jumlah P yang diinput adalah 1. Terdapat 351 penjualan pada bulan Desember 2021, oleh karena itu sistem akan mencoba mencari bulan dengan jarak yang paling mendekati angka 351 menggunakan metode *Euclidean Distance*. Hasil perhitungan jarak menunjukkan bahwa bulan Juni 2020 dengan penjualan 346 dan Oktober 2020 dengan penjualan 334 merupakan bulan dengan jarak terpendek. Hasil prediksi untuk bulan Januari 2022 adalah melakukan perhitungan rata-rata terhadap penjualan bulan Juli 2020 dan November 2020 seperti pada persamaan (7).

$$\text{Rata - rata} = \frac{198+328}{2} \quad (7)$$

$$\text{Rata - rata} = 263 \quad (8)$$

Perhitungan akurasi menggunakan metode *RMSE* adalah seperti persamaan (9), (10) dan (11).

$$RMSE = \sqrt{\frac{\sum (x_i - y_i)^2}{n}} \quad (9)$$

$$RMSE = \sqrt{\frac{\sum (502-263)^2 + (436-518)^2}{2}} \quad (10)$$

$$RMSE = 178,6 \quad (11)$$

Perhitungan akurasi menggunakan metode *MAPE* adalah seperti persamaan (12), (13), dan (14).

$$MAPE = \frac{1}{n} \sum \left| \frac{At - Ft}{At} \right| \quad (12)$$

$$MAPE = \frac{1}{2} \sum \left| \frac{502-263}{502} \right| + \left| \frac{436-518}{436} \right| \quad (13)$$

$$MAPE = 0,332 \quad (14)$$

Tabel 2 merupakan hasil uji coba untuk menghitung akurasi prediksi menggunakan metode *RMSE* dan *MAPE* untuk prediksi berdasarkan *product* menggunakan data penjualan 10 *product*

terlaris. *Input* parameter yang digunakan adalah *K* senilai 2 dan *P* senilai 1.

Tabel 2. Hasil Akurasi Prediksi *Product* dengan *K*=1 dan *P*=1 (*KNN Regresi*)

Product	K=2,P=1	
	RMSE	MAPE
INDOMIE GORENG/40	178,66	0,33
MASAKO AYAM 11 G (1`S)/720	295,70	1,39
AQUA BOTOL 600 ML/24	236,95	0,44
BEAR BRAND 189 ML/30	277,04	0,34
TEH PUCUK HARUM JASMINE 350 ML/24	129,17	0,27
GULA PASIR LOKAL 950GR	74,91	0,24
AQUA BOTOL 330 ML/24	77,91	0,35
AQUA BOTOL 1500 ML/12	50,08	0,16
YAKULT 5`S	26,80	0,18
GUDANG GARAM SURYA 12	247,14	0,66
Rata - Rata	159,44	0,44

Selanjutnya uji coba akan dilakukan dengan mengubah *input* parameter menjadi kombinasi antara *K*=1 sampai *K*=3 dan *P*=1 sampai *P*=3. Hasil akurasi dari uji coba ini menunjukkan bahwa *input* parameter *K*=3 dan *P*=1 akan menghasilkan tingkat akurasi terbaik dengan rata-rata *RMSE*= 112,93 dan *MAPE*= 0,27 yang dapat dilihat dari Tabel 3.

Tabel 3. Hasil Akurasi Prediksi *Product* (*KNN Regresi*)

Product	K=3,P=1	
	RMSE	MAPE
INDOMIE GORENG/40	150,76	0,32
MASAKO AYAM 11 G (1`S)/720	92,03	0,42
AQUA BOTOL 600 ML/24	251,04	0,40
BEAR BRAND 189 ML/30	227,28	0,30
TEH PUCUK HARUM JASMINE 350 ML/24	45,52	0,12
GULA PASIR LOKAL 950GR	69,99	0,24
AQUA BOTOL 330 ML/24	89,32	0,38
AQUA BOTOL 1500 ML/12	23,21	0,06
YAKULT 5`S	32,22	0,20
GUDANG GARAM SURYA 12	147,92	0,27
Rata - Rata	112,93	0,27

5.2. Pengujian Nilai P dan K untuk KNN Regresi (Divisi)

Selanjutnya adalah uji coba prediksi berdasarkan divisi menggunakan 8 divisi yang ada di Supermarket X. Berdasarkan uji coba diatas, maka didapatkan bahwa prediksi berdasarkan divisi dengan metode *KNN Regresi* menggunakan *input* parameter *K*=2 dan *P*=2 akan menghasilkan akurasi terbaik dengan rata-rata *RMSE*= 1728,11 dan *MAPE*= 0,45 seperti pada Tabel 4.

Tabel 4. Hasil Akurasi Prediksi Divisi (*KNN Regresi*)

Divisi	K=2,P=2	
	RMSE	MAPE
GROCERY FOOD	8.125,69	0,36
GROCERY NON FOOD	3.874,12	0,49
FRESH PRODUCT	283,77	0,24
GENERAL MERCHANDISE	432,20	0,41
FASHION/TEXTILES DEWASA	62,22	0,77
FASHION/TEXTILES BABY	46,67	0,28
FASHION	17,15	0,89
OPERASIONAL	983,06	0,15
Rata - Rata	1.728,11	0,45

5.3. Pengujian Nilai P dan K untuk KNN Regresi (Department)

Selanjutnya adalah uji coba prediksi berdasarkan *department* menggunakan 10 *department* yang ada di Supermarket X. Hasil akurasi dari uji coba ini menunjukkan bahwa *input* parameter K=3 dan P=1 akan menghasilkan tingkat akurasi terbaik dengan rata-rata RMSE= 1063,32 dan MAPE= 0,41 yang dapat dilihat dari Tabel 5.

Tabel 5. Hasil Akurasi Prediksi Departemen (KNN Regresi)

Department	K=3,P=1	
	RMSE	MAPE
BEVERAGES	2.088,45	0,23
FOOD2	2.960,35	0,29
FOOD	2.805,34	0,31
PERSONAL CARE	1.518,45	0,30
PERSONAL CARE 2	249,68	0,16
PERISHABLE	377,64	0,27
FRUIT	59,60	0,85
HOUSE HOLD	259,51	0,56
HOME CARE	289,80	0,17
STATIONERY	24,34	0,93
Rata - Rata	1.063,32	0,41

5.4. Pengujian Nilai P dan K untuk KNN Regresi (Kategori)

Selanjutnya adalah uji coba prediksi berdasarkan kategori menggunakan 10 kategori yang ada di Supermarket X. Hasil akurasi dari uji coba ini menunjukkan bahwa *input* parameter K=3 dan P=1 akan menghasilkan tingkat akurasi terbaik dengan rata-rata RMSE= 604,50 dan MAPE= 0,35 yang dapat dilihat dari Tabel 6.

Tabel 6. Hasil Akurasi Prediksi Kategori (KNN Regresi)

Kategori	K=3,P=1	
	RMSE	MAPE
BREAKFAST	1.556,14	0,57
BODY CARE	466,21	0,38
BUMBU EASTERN	181,50	0,15
AIR MINERAL	890,79	0,35
BISKUIT	743,81	0,32
SUSU CAIR	562,72	0,26
SUSU	514,81	0,46
MIE	194,96	0,08
HAIR CARE	697,47	0,73
PERAWATAN PAKAIAN	236,62	0,23
Rata - Rata	604,50	0,35

5.5. Pengujian Nilai P dan K untuk KNN Regresi (Sub Kategori)

Selanjutnya adalah uji coba prediksi berdasarkan sub kategori menggunakan 10 sub kategori yang ada di Supermarket X. Hasil akurasi dari uji coba ini menunjukkan bahwa *input* parameter K=3 dan P=1 akan menghasilkan tingkat akurasi terbaik dengan rata-rata RMSE= 280 dan MAPE= 0,30 yang dapat dilihat dari Tabel 7

Tabel 7. Hasil Akurasi Prediksi Sub Kategori (KNN Regresi)

Sub Kategori	K=3,P=1	
	RMSE	MAPE
GREEN TEA	175,79	0,25
BREAD/ROTI	113,86	0,13
ES KRIM	219,41	0,42
KOPI INSTANT	214,29	0,22
ROKOK FILTER	695,66	0,51
AIR 600 ML	419,44	0,32
MIE INSTANT	242,78	0,08

SUSU CAIR KOTAK	261,53	0,32
SHAMPO DEWASA	331,99	0,43
PEWANGI PAKAIAN	125,27	0,32
Rata - Rata	280,00	0,30

5.6. Pengujian Nilai Season untuk Multiplicative Decomposition (Produk)

Metode *multiplicative decomposition* digunakan untuk memprediksi data yang mempunyai pola musiman dan trend. Data akan dibagi menjadi beberapa *season* atau musim dan setiap bulan akan dicari *seasonal index*nya. Input jumlah *season* akan berpengaruh terhadap tingkat akurasi prediksi sehingga pada penelitian ini, akan dilakukan uji coba untuk mengetahui berapa nilai *season* yang memiliki tingkat akurasi tertinggi untuk setiap prediksi terhadap produk, *departemen*, divisi, kategori, dan sub kategori. Uji coba akan dilakukan dengan menggunakan nilai *season* sebanyak 1 sampai 6. Tabel 1 diatas merupakan daftar transaksi yang terjadi setiap bulannya pada *product* “INDOMIE GORENG/40”. Berdasarkan uji coba, maka didapatkan bahwa prediksi berdasarkan *product* dengan metode *Multiplicative Decomposition* menggunakan *input* parameter *Season*= 4 akan menghasilkan akurasi terbaik dengan rata-rata RMSE= 122,52 dan MAPE= 0,27. Tabel 8 adalah hasilnya.

Tabel 8. Hasil Akurasi Prediksi Produk (Multiplicative Decomposition)

Product	Season=4	
	RMSE	MAPE
INDOMIE GORENG/40	42,84	0,08
MASAKO AYAM 11 G (1'S)/720	62,36	0,30
AQUA BOTOL 600 ML/24	195,91	0,32
BEAR BRAND 189 ML/30	257,78	0,34
TEH PUCUK HARUM JASMINE 350 ML/24	122,31	0,32
GULA PASIR LOKAL 950GR	100,93	0,24
AQUA BOTOL 330 ML/24	108,36	0,20
AQUA BOTOL 1500 ML/12	121,87	0,37
YAKULT 5'S	18,56	0,10
GUDANG GARAM SURYA 12	194,24	0,43
Rata - Rata	122,52	0,27

5.7. Pengujian Nilai Season untuk Multiplicative Decomposition (Divisi)

Selanjutnya adalah uji coba prediksi berdasarkan divisi menggunakan 8 divisi yang ada di Supermarket X. Berdasarkan uji coba, maka didapatkan bahwa prediksi berdasarkan divisi dengan metode *Multiplicative Decomposition* menggunakan *input* parameter *Season*=1 akan menghasilkan akurasi terbaik dengan rata-rata RMSE= 1108,58 dan MAPE= 0,41. Tabel 9 adalah hasil akurasi.

Tabel 9. Hasil Akurasi Prediksi Divisi (Multiplicative Decomposition)

Divisi	Season=1	
	RMSE	MAPE
GROCERY FOOD	4.804,86	0,21
GROCERY NON FOOD	652,89	0,06
FRESH PRODUCT	217,9	0,13
GENERAL MERCHANDISE	444,23	0,41
FASHION/TEXTILES DEWASA	68,42	0,86
FASHION/TEXTILES BABY	164,94	0,82
FASHION	13,36	0,42
OPERASIONAL	2.502,00	0,4
Rata - Rata	1.108,58	0,41

5.8. Pengujian Nilai Season untuk Multiplicative Decomposition (Departemen)

Selanjutnya adalah uji coba prediksi berdasarkan *department* menggunakan 10 *department* yang ada di Supermarket X. Berdasarkan uji coba, maka didapatkan bahwa prediksi berdasarkan *department* dengan metode *Multiplicative Decomposition* menggunakan *input* parameter *Season*= 4 akan menghasilkan akurasi terbaik dengan rata-rata *RMSE*=693,37 dan *MAPE*= 0,39. Tabel 10 adalah hasil akurasinya.

Tabel 10. Hasil Akurasi Prediksi Departemen (Multiplicative Decomposition)

Department	Season= 4	
	RMSE	MAPE
BEVERAGES	1.862,85	0,22
FOOD2	1.147,71	0,1
FOOD	2.120,63	0,21
PERSONAL CARE	104,69	0,01
PERSONAL CARE 2	533,21	0,3
PERISHABLE	268,36	0,27
FRUIT	147,4	1,7
HOUSE HOLD	230,56	0,49
HOME CARE	434,11	0,2
STATIONERY	84,22	0,4
Rata - Rata	693,37	0,39

5.9. Pengujian Nilai Season untuk Multiplicative Decomposition (Kategori)

Selanjutnya adalah uji coba prediksi berdasarkan kategori menggunakan 10 kategori yang ada di Supermarket X. Berdasarkan uji coba, maka didapatkan bahwa prediksi berdasarkan kategori dengan metode *Multiplicative Decomposition* menggunakan *input* parameter *Season*= 4 akan menghasilkan akurasi terbaik dengan rata-rata *RMSE*= 341,89 dan *MAPE*= 0,17. Tabel 11 adalah hasil akurasinya.

Tabel 11. Hasil Akurasi Prediksi Kategori (Multiplicative Decomposition)

Kategori	Season= 4	
	RMSE	MAPE
BREAKFAST	288,39	0,07
BODY CARE	307,36	0,22
BUMBU EASTERN	53,40	0,04
AIR MINERAL	848,36	0,32
BISKUIT	440,29	0,18
SUSU CAIR	570,97	0,23
SUSU	346,87	0,24
MIE	200,24	0,08
HAIR CARE	129,26	0,13
PERAWATAN PAKAIAN	233,77	0,2
Rata - Rata	341,89	0,17

5.10. Pengujian Nilai Season untuk Multiplicative Decomposition (Sub Kategori)

Selanjutnya adalah uji coba prediksi berdasarkan sub kategori menggunakan 10 sub kategori yang ada di Supermarket X. Berdasarkan uji coba, maka didapatkan bahwa prediksi berdasarkan subkategori dengan metode *Multiplicative Decomposition* menggunakan *input* parameter *Season*= 4 akan menghasilkan akurasi terbaik dengan rata-rata *RMSE*= 198,10 dan *MAPE*= 0,20. Prediksi untuk sub kategori “Es Krim” memberikan *RMSE* terbaik yaitu 23,50 sedangkan sub kategori “Roti” memberikan hasil *MAPE* terbaik yaitu 0,02. Tabel 12 adalah hasil akurasinya.

Tabel 12. Hasil Akurasi Prediksi Sub Kategori (Multiplicative Decomposition)

Sub Kategori	Season= 4	
	RMSE	MAPE
GREEN TEA	209,44	0,25
BREAD/ROTI	34,26	0,02
ES KRIM	23,50	0,04
KOPI INSTANT	265,74	0,20
ROKOK FILTER	640,73	0,42
AIR 600 ML	332,66	0,35
MIE INSTANT	131,57	0,04
SUSU CAIR KOTAK	61,90	0,06
SHAMPO DEWASA	54,37	0,08
PEWANGI PAKAIAN	226,82	0,5
Rata - Rata	198,10	0,20

5.11. Perbandingan Akurasi Antar Metode

Setelah mendapatkan nilai *input* parameter yang optimal dari kedua metode, maka selanjutnya akan dilakukan perhitungan terhadap nilai *RMSE* dan *MAPE* dari kedua metode dengan *input* parameter yang memberikan hasil akurasi terbaik. Tabel 13 merupakan perbandingan dari hasil perhitungan akurasi terbaik kedua metode menggunakan *RMSE* dan *MAPE* berdasarkan uji coba yang dilakukan pada bab-bab diatas. Perbandingan akan dilakukan baik untuk prediksi berdasarkan divisi, *department*, kategori, sub kategori, maupun *product*.

Tabel 13. Perbandingan Akurasi

	KNN Regresi		Multiplicative Decomposition	
	RMSE	MAPE	RMSE	MAPE
Divisi	1.728,11	0,45	1.108,58	0,41
Department	1.063,32	0,41	693,37	0,39
Kategori	604,50	0,35	341,89	0,17
Sub Kategori	280,00	0,30	198,10	0,20
Product	112,93	0,27	122,52	0,27
Rata-Rata	757,77	0,36	492,89	0,29

6. KESIMPULAN

Program dapat menampilkan hasil *forecasting* berdasarkan 2 metode yaitu *KNN Regresi* dan *Multiplicative Decomposition*. Metode yang paling tepat digunakan untuk prediksi penjualan pada Supermarket X adalah metode *Multiplicative Decomposition* karena memberikan hasil akurasi yang lebih baik daripada *KNN Regresi*. Prediksi berdasarkan divisi memberikan tingkat akurasi terbaik bila menggunakan *input* parameter *K*=2 dan *P*=2 untuk metode *KNN Regresi* sebesar 1728,11 (*RMSE*) dan 0,45 (*MAPE*), sedangkan untuk metode *Multiplicative Decomposition* menggunakan *input* parameter *Season*= 1 sebesar 1108,58 (*RMSE*) dan 0,41 (*MAPE*). Prediksi berdasarkan *department* memberikan tingkat akurasi terbaik bila menggunakan *input* parameter *K*=3 dan *P*=1 untuk metode *KNN Regresi* sebesar 1063,32 (*RMSE*) dan 0,41 (*MAPE*), sedangkan untuk metode *Multiplicative Decomposition* menggunakan *input* parameter *Season*= 4 sebesar 693,37 (*RMSE*) dan 0,39 (*MAPE*). Prediksi berdasarkan kategori memberikan tingkat akurasi terbaik bila menggunakan *input* parameter *K*=3 dan *P*=1 untuk metode *KNN Regresi* sebesar 604,50 (*RMSE*) dan 0,35 (*MAPE*), sedangkan untuk metode *Multiplicative Decomposition* menggunakan *input* parameter *Season*= 4 sebesar 341,89 (*RMSE*) dan 0,17 (*MAPE*). Prediksi berdasarkan sub kategori memberikan tingkat akurasi terbaik bila menggunakan *input* parameter *K*=3 dan *P*=1 untuk metode *KNN Regresi* sebesar 280,0 (*RMSE*) dan 0,30 (*MAPE*), sedangkan untuk metode *Multiplicative Decomposition* menggunakan *input* parameter *Season*= 4 sebesar 198,10 (*RMSE*) dan 0,20 (*MAPE*). Prediksi berdasarkan produk memberikan

tingkat akurasi terbaik bila menggunakan *input* parameter $K=3$ dan $P=1$ untuk metode *KNN Regresi* sebesar 112,93 (RMSE) dan 0,27 (MAPE), sedangkan untuk metode *Multiplicative Decomposition* menggunakan *input* parameter *Season*= 4 sebesar 122,52 (RMSE) dan 0,27 (MAPE).

7. SARAN

Berikut adalah beberapa saran yang dapat dilakukan untuk mengembangkan penelitian lebih lanjut:

- Mengembangkan desain dan tampilan program agar lebih mudah digunakan.
- Melakukan integrasi dengan sistem perusahaan agar *user* tidak perlu melakukan *import* dataset lagi.

8. REFERENSI

- [1] Ayuni, G. N., & Fitriana, D. (2019). Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ. *Jurnal Telematika*, 14(2), 79-86.
- [2] Azis, F., Defiyanti, S., & Sari, B. N. (2018). Perbandingan Algoritma Cart Dan *K-Nearest Neighbor*. *Technology Acceptance Model*, 9(2), 74-78.
- [3] Budi, E., Santoso, L. W., & Dewi, L. P. (2019). Perancangan Dan Pembuatan *Data Warehouse* Dan *Business Intelligence* Pada *Market Research Motorcycle* Honda Mpm Motor. *Jurnal Infra Universitas Kristen Petra*, 7(2), 88-94.
- [4] Bustomy, M. I. (2020). Implementasi *Business Intelligence* untuk Prestasi Mahasiswa STTI NIIT. *Jurnal JI-Tech*, 16(1), 1-11.
- [5] Cristian, M. (2018). *Average Monthly Rainfall Forecast in Romania by Using K-Nearest Neighbors Regression*. *Annals - Economy Series*, (4).
- [6] Darmayanti, I., Subarkah, P., Anunggilarso, L. R., & Suhaman, J. (2021). Prediksi Potensi Siswa Putus Sekolah Akibat Pandemi Covid-19 Menggunakan Algoritme *K-Nearest Neighbor*. *Jurnal Sains dan Teknologi*, 10(2). DOI=10.23887/jstundiksha.v10i2.39151.
- [7] Hasmawati, Nangi, J., & Muchtar, M. (2017). Aplikasi Prediksi Penjualan Barang Menggunakan Metode *K-Nearest Neighbor(Knn)* (Studi Kasus Tumaka Mart). *semanTIK*.
- [8] Hijriani, A., Aprilliana, E., Pribadi, R. I., & Sakethi, D. (2020). *Business Intelligence Dashboard (BID)* Pada Usaha Mikro Bidang Retail Studi Kasus CV Duta Square Bandar Lampung. *Jurnal Manajemen dan Teknologi Informatika*, 10(1), 11-18. DOI=10.31940/matrix.v10i1.1616.
- [9] Indarwati, T., Irawati, T., & Rimawati, E. (2018). Penggunaan Metode *Linear Regression* Untuk Prediksi Penjualan *Smartphone*. *Jurnal Teknologi Informasi dan Komunikasi Sinar Nusantara*, 6(2). DOI=10.30646/tikomsin.v6i2.369.
- [10] Kristiyanti, D. A., & Sumarno, Y. (2020). Penerapan Metode *Multiplicative Decomposition (Seasonal)* Untuk Peramalan Persediaan Barang Pada PT. Agrinusa Jaya Santosa. *Jurnal Sistem Komputer dan Kecerdasan Buatan*, 3(2), 45-51.
- [11] Nanja, M., & Purwanto. (2015). Metode *K-Nearest Neighbor* Berbasis *orward Selection* Untuk Prediksi Harga Komoditi Lada. *Jurnal Pseudocode*, 2(1), 53-64. DOI=10.33369/pseudocode.2.1.53-64.
- [12] Saputra, I., Alkadri, S., & Insani, R. W. (2021). Sistem Pendukung Keputusan Pemilihan Penerima Beasiswa Universitas Muhammadiyah Pontianak Menggunakan Metode *Fuzzy Mamdani*. *Digital Intelligence*, 2(1). DOI=10.29406/diligent.v2i1.2903.
- [13] Seruni, D., Furqon, M., & Wihandika, R. (2020). Sistem Prediksi Pertumbuhan Jumlah Penduduk Kota Malang menggunakan Metode *K-Nearest Neighbor Regression*. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 4(4), 1075-1082.
- [14] Supardi, R. (2020). Penerapan Metode *Regresi Linear* Dalam Memprediksi Data Penjualan Barang Di Toko Bangunan Vita Viya. *Journal of Technopreneurship and Information System (JTIS)*, 3(1), 11-18. DOI=10.36085/jtis.v3i1.629.
- [15] Syaifulloh, A. (2018). Perbandingan 6 Metode *Forecasting* Dalam Peramalan Jumlah Maba Stmik Ppkia Pradnya Paramita Malang. *Teknologi Informasi: Teori, Konsep, dan Implementasi: Jurnal Ilmiah*, 9(2), 91-98.
- [16] Tanuwijaya, J., & Hansun, S. (2019). *LQ45 Stock Index Prediction using k-Nearest Neighbors Regression*. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(3), 2388 - 2391. DOI=10.35940/ijrte.C4663.098319.