

Penerapan Algoritma Regresi Linear untuk Prediksi Penjualan Beberapa Produk Makanan dan Minuman (Studi Kasus: Kantin Berkat)

Ribkha Angela Tiurma¹⁾, Ridha Sefina Samosir²⁾

^{1,2)} Sistem Informasi, Fakultas Ilmu Komputer dan Desain, Universitas Kalbis
Jalan Pulomas Selatan Kav. 22, Jakarta 13210

¹⁾ Email: 2019104774@student.kalbis.ac.id

²⁾ Email: ridha.samosir@kalbis.ac.id

Abstract: Sales are activities that provides benefits for business operators, including UMKM. Every business operator has documentation of sales data. Kantin Berkat, UMKM which sells food and beverages, always documents their daily sales. Documentation is to record the daily sales of products. Daily sales of products in Kantin Berkat vary that some products show sales fluctuations. The impact is some products must be wasted because they don't sell on certain days and must have affected the gain at Kantin Berkat. This research attempts to predict the sales amount for the next days based on sales data of two years prior. The prediction's done with the Simple Linear Regression Algorithm by Python encoding for five products, like seblak, dimsum, singkong, es jeruk, and es teh manis. The results of testing the algorithm showed that the MAPE values for the prediction results of the five products were 35,86%, 28,17%, 48,20%, 15,92%, and 20,03%.

Keywords: MAPE, prediction, python, sales data, simple linear regression

Abstrak: Penjualan merupakan aktivitas yang memberikan keuntungan bagi para pelaku usaha, termasuk Usaha Mikro Kecil dan Menengah. Setiap pelaku usaha memiliki dokumentasi data penjualannya. Kantin Berkat, UMKM yang menjual produk makanan dan minuman, selalu mendokumentasikan jumlah penjualan hariannya. Dokumentasi dilakukan untuk mencatat jumlah penjualan produk setiap hari. Penjualan produk di Kantin Berkat bervariasi, bahkan pada beberapa produk menunjukkan fluktuasi jumlah penjualan. Dampak dari kondisi ini adalah sering terjadinya beberapa produk yang harus terbuang karena pada hari tertentu produk tersebut tidak laku sehingga mempengaruhi perolehan keuntungan di Kantin Berkat. Penelitian mencoba untuk melakukan prediksi jumlah penjualan pada hari berikutnya berdasarkan data penjualan yang terjadi dua tahun sebelumnya. Prediksi dilakukan dengan Algoritma Regresi Linear Sederhana dengan pengkodean Python untuk lima produk, yaitu seblak, dimsum, singkong, es jeruk, dan es teh manis. Hasil pengujian terhadap algoritma menunjukkan bahwa nilai MAPE untuk hasil prediksi kelima produk secara berturut-turut adalah 35,86%, 28,17%, 48,20%, 15,92%, dan 20,03%.

Kata kunci: MAPE, prediksi, python, data penjualan. regresi linear sederhana

I. PENDAHULUAN

Penjualan merupakan salah satu hal yang vital bagi suatu perusahaan atau organisasi karena dengan melakukan kegiatan penjualan, baik barang maupun jasa, perusahaan atau organisasi akan mendapatkan keuntungan. Penjualan tidak hanya dilakukan oleh perusahaan atau organisasi besar, tetapi juga Usaha Mikro Kecil dan Menengah (UMKM). UMKM merupakan usaha dengan skala yang kecil dan dapat meliputi usaha atau bisnis apapun, termasuk usaha atau bisnis kuliner. Tidak jarang ditemukan UMKM berlomba-lomba untuk mendapatkan banyak keuntungan dari kegiatan penjualannya, apalagi dengan usaha atau bisnis yang serupa. Hal ini

pula didukung oleh era globalisasi yang mana teknologi sudah berkembang pesat sehingga kegiatan penjualan dapat dilakukan lebih efektif dan efisien. Segala sesuatunya dapat dilakukan secara daring, termasuk penjualan produknya.

Oleh sebab itu, setiap pemilik usaha perlu menyusun strategi bisnisnya agar tetap mendapatkan keuntungan dari kegiatan penjualannya tersebut di tengah banyaknya bisnis serupa, yakni kuliner, yang semakin bertumbuh dan berkembang. Penyusunan strategi bisnis ini dapat diawali dengan melakukan prediksi, perkiraan, atau peramalan penjualan produk yang disediakan pada masa mendatang.

Prediksi atau peramalan merupakan proses

memperkirakan dan memperhitungkan secara sistematis mengenai keadaan yang paling mungkin terjadi di masa depan berdasarkan informasi dan pengujian di masa lalu. Prediksi yang dilakukan ini tidak harus memberikan jawaban yang pasti, tetapi berusaha untuk mendapatkan jawaban sedekat mungkin yang akan terjadi [1]. Prediksi penjualan biasanya dilakukan untuk suatu tujuan, yaitu membantu pemilik usaha mengambil keputusan dalam menyediakan produk yang dijualnya. Prediksi penjualan yang dilakukan diharapkan mampu memberikan hasil yang baik sehingga risiko kesalahan dapat ditekan seminimal mungkin.

Kantin Berkat merupakan UMKM yang bergerak di bidang kuliner yang menjual makanan rumahan dan minuman. Makanan yang disediakan oleh Kantin Berkat beragam, bahkan dapat berbeda setiap harinya. Namun, Kantin Berkat tentu memiliki makanan dan minuman tetap, yang selalu sama setiap harinya. Penjualan setiap hari keseluruhan makanan dan minuman dalam satu bulan kira-kira mencapai 200 porsi. Meskipun demikian, Kantin Berkat tetap mengalami permasalahan selama kegiatan penjualannya. Permasalahan utama yang dialami oleh Kantin Berkat adalah permintaan pembeli akan menu yang disediakan tidak menentu sehingga jumlah penjualan produk makanan dan minuman cenderung fluktuatif dan tidak dapat ditentukan secara gamblang. Fluktuasi tersebut menyebabkan banyak stok produk makanan dan minuman yang sudah disediakan menjadi tidak laku dan berakhir harus dibuang.

Berdasarkan latar belakang tersebut, diusulkan untuk dilakukan suatu prediksi penjualan guna mempermudah pemilik Kantin Berkat untuk mengetahui jumlah penjualan di waktu selanjutnya dari beberapa produk penjualannya. Dengan demikian, dapat diketahui perkiraan stok yang perlu disediakan dan mampu meminimalkan produk yang tidak laku dan harus dibuang. Prediksi jumlah penjualan dilakukan untuk beberapa produk makanan dan minuman Kantin Berkat, yakni seblak, dimsum, singkong, es jeruk, dan es teh manis, yang didasarkan pada pernyataan pemilik.

Berdasarkan latar belakang tersebut, diusulkan untuk dilakukan suatu prediksi penjualan guna mempermudah pemilik Kantin Berkat untuk mengetahui jumlah penjualan di waktu

selanjutnya dari beberapa produk penjualannya. Dengan demikian, dapat diketahui perkiraan stok yang perlu disediakan dan mampu meminimalkan produk yang tidak laku dan harus dibuang. Prediksi jumlah penjualan dilakukan untuk beberapa produk makanan dan minuman Kantin Berkat, yakni seblak, dimsum, singkong, es jeruk, dan es teh manis, yang didasarkan pada pernyataan pemilik bahwa kelima produk tersebut merupakan produk yang lebih laku dan penjualannya lebih fluktuasi dibandingkan produk lainnya. Prediksi jumlah penjualan difokuskan pada penerapan algoritma *machine learning*, yaitu Algoritma Regresi Linear Sederhana dengan batasan masalah, yaitu data yang digunakan merupakan data historis penjualan produk makanan dan minuman dari tahun 2021 sampai dengan 2022, prediksi penjualan yang dilakukan untuk setiap hari dimulai dari bulan Juni sampai dengan Desember 2023, dan data prediksi penjualan dilakukan untuk lima produk Kantin Berkat. Tujuan dari penelitian ini adalah untuk menerapkan Algoritma Regresi Linear Sederhana dalam memprediksi jumlah penjualan beberapa produk Kantin Berkat di masa mendatang.

II. METODE PENELITIAN

Bagian ini berisi penjabaran beberapa teori pendukung penelitian dan metodologi penelitian yang digunakan sebagai berikut.

Secara umum, data terbagi menjadi dua jenis, yaitu data internal dan data eksternal. Data internal merupakan data asli yang memaparkan kondisi suatu perusahaan atau organisasi yang dihasilkan melalui pengamatan dan penelitian dari pihak internal perusahaan atau organisasi tersebut. Data eksternal merupakan data yang memaparkan kondisi di luar perusahaan atau organisasi yang dihasilkan melalui pengamatan atau penelitian pihak eksternal perusahaan atau organisasi tersebut [2], [3]. Salah satu contoh data internal yang sederhana adalah data penjualan.

Dalam dunia bisnis, kegiatan penjualan dikenal sebagai salah satu faktor yang mampu untuk mengembangkan suatu usaha, baik menjual produk barang maupun jasa. Suatu usaha akan berkembang apabila terdapat keterlibatan dua pihak, yaitu penjual dan pembeli. Dengan demikian, penjualan dapat dijelaskan sebagai suatu kegiatan yang dilakukan penjual

untuk mencari pembeli, mempengaruhi dan mengarahkan pembeli untuk menyesuaikan kebutuhannya dengan produk yang ditawarkan, serta mengadakan kesepakatan terkait harga yang memberikan keuntungan kepada dua belah pihak, yaitu penjual dan pembeli [4]. Kegiatan penjualan, dalam kehidupan sehari-hari, tidak hanya dilakukan oleh perusahaan atau organisasi besar, tetapi juga dilakukan oleh individu secara umum. Penjualan tersebut tergolong ke dalam penjualan dalam skala kecil yang disebut sebagai Usaha Mikro Kecil dan Menengah (UMKM).

Usaha Mikro Kecil dan Menengah atau yang disingkat sebagai UMKM merupakan usaha atau bisnis yang dimiliki dan dilakukan oleh perorangan ataupun badan usaha, yang telah memenuhi kriteria sebagai usaha atau bisnis mikro [5]. Menurut Undang- Undang No. 20 Tahun 2008, kriteria UMKM berdasarkan kekayaan bersih (tidak termasuk tanah dan bangunan tempat usaha) dapat dijabarkan sebagai berikut [6].

1. Usaha Mikro, dengan kekayaan bersih paling banyak Rp50.000.000,- atau hasil penjualan tahunan paling banyak Rp300.000.000,-.
2. Usaha Kecil, dengan kekayaan bersih lebih dari Rp50.000.000,- sampai dengan paling banyak Rp500.000.000,- atau hasil penjualan tahunan lebih dari Rp300.000.000,- sampai dengan paling banyak Rp2.500.000.000,-.
3. Usaha Menengah, dengan kekayaan bersih lebih dari Rp300.000.000,- sampai dengan Rp10.000.000.000,- atau hasil penjualan tahunan lebih dari Rp2.500.000.000,-

Sebagai suatu usaha atau bisnis yang produktif dan semakin berkembang, UMKM memiliki keterlibatan yang tinggi dalam meningkatkan kesejahteraan rakyat sehingga mampu mendorong pergerakan pembangunan dan perekonomian Indonesia [7]. Salah satu bidang UMKM yang banyak ditemui dan mampu menjadi peluang bagi banyak individu serta semakin berkembang adalah bisnis kuliner [8]. Hal ini disebabkan makanan merupakan kebutuhan pokok semua orang dan paling banyak dicari. Salah satu contoh bisnis kuliner sebagai bagian dari UMKM adalah

kantin.

Prediksi atau peramalan atau *forecasting* dapat diasumsikan sebagai kegiatan untuk memperkirakan kondisi yang akan terjadi di masa mendatang [9]. Kegiatan ini dilakukan dengan menggunakan data masa lalu yang didokumentasikan dari sebuah variabel atau sekumpulan variabel [10]. Dengan kata lain, prediksi atau peramalan merupakan proses memperkirakan dan memperhitungkan, secara sistematis, keadaan yang paling mungkin terjadi di masa depan berdasarkan informasi atau pengujian yang dilakukan sebelumnya

Prediksi atau peramalan pada suatu perusahaan atau organisasi banyak dilakukan untuk memberikan gambaran mengenai keadaan perusahaan atau organisasi di masa depan yang mana gambaran tersebut dapat digunakan untuk membantuk pemilik mengambil keputusan dalam rangka memenuhi permintaan pembeli [11]. Prediksi yang dilakukan tidaklah harus memberikan jawaban yang pasti mengenai kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin mengenai kejadian tersebut [12]. Salah satu manfaat prediksi yang dapat dilakukan adalah prediksi penjualan.

Prediksi penjualan merupakan proses yang dilakukan oleh perusahaan atau organisasi untuk memperkirakan pendapatan dalam beberapa waktu ke depan berdasarkan perkiraan jumlah produk yang terjual [13]. Prediksi penjualan ini juga mampu memperkirakan besarnya penjualan produk pada waktu tertentu. Dengan demikian, prediksi penjualan ini menjadi salah satu hal terpenting dalam menyusun rencana produksi suatu perusahaan atau organisasi untuk mengendalikan jumlah stok produk yang perlu disediakan sehingga produk yang tidak laku dan harus dibuang dapat diminimalkan.

A. Objek Penelitian

Objek penelitian yang menjadi fokus penelitian ini adalah data historis penjualan beberapa produk makanan dan minuman milik Kantin Berkat, yakni seblak, dimsum, singkong, es jeruk, dan es teh manis, dari tahun

2021 sampai dengan 2022. Data tersebut terdiri dari dua variabel atau atribut data, yaitu atribut tanggal dengan tipe data *date* (dd/mm/yyyy) dan atribut jumlah penjualan dengan tipe data *integer*.

B. Lokasi dan Waktu Penelitian

Penelitian ini dilakukan di Kantin Berkat yang terletak pada kawasan kuliner, yaitu VJS Kuliner di Jalan Sedap Malam, Jaka Setia, Bekasi Selatan, Kota Bekasi, Jawa Barat.

Penelitian ini dilakukan sejak bulan November 2022 untuk permohonan dan perizinan mitra, Januari 2023 untuk pengumpulan data, Maret 2023 untuk penerimaan data penjualan, sampai pada penelitian ini selesai dilaksanakan pada bulan Juni 2023.

C. Metode Penelitian

Metode penelitian yang digunakan pada penelitian ini adalah metode penelitian kualitatif yang cenderung kepada suatu analisis deskriptif. Metode pengumpulan data yang dilakukan, yaitu (1) Observasi, dengan melakukan pengamatan secara langsung di Kantin Berkat untuk mengetahui kondisi sebenarnya yang terjadi; (2) Wawancara, dengan mengajukan pertanyaan sederhana kepada pemilik Kantin Berkat mengenai latar belakang, kondisi, dan permasalahan yang dihadapi Kantin Berkat selama beroperasi; (3) Dokumentasi, dengan mengumpulkan data yang sudah ada dan sudah didokumentasikan sebelumnya oleh pemilik Kantin Berkat, yakni data historis penjualan beberapa produk untuk dua tahun terakhir; dan (4) Studi literatur, dengan mengumpulkan informasi-informasi pendukung yang relevan dengan penelitian ini dari berbagai sumber, seperti jurnal dan e-book.

Algoritma yang digunakan dalam penelitian ini adalah Algoritma Regresi Linear Sederhana. Algoritma regresi merupakan teknik machine learning yang digunakan untuk memprediksi nilai dari sebuah variabel target berdasarkan pada nilai dari beberapa variabel input yang sudah ada sebelumnya. Regresi biasanya digunakan untuk menentukan hubungan sebab-akibat (linier) antara satu

variabel terikat dengan variabel bebas lainnya. Selain itu, regresi seringkali juga digunakan untuk menilai kekuatan hubungan antarvariabel dalam memprediksi suatu keputusan di masa depan yang mampu meminimalkan kesalahan dalam pengambilan keputusan dan menjadi bahan evaluasi suatu perusahaan atau organisasi [14], [15].

Regresi linear merupakan tipe regresi yang paling umum, mudah diimplementasikan, dan sederhana. Regresi linear terbagi menjadi dua, yaitu Regresi Linear Sederhana dan Regresi Linear Berganda. Regresi Linear Sederhana hanya melibatkan satu variabel bebas, sedangkan Regresi Linear Berganda melibatkan lebih dari satu variabel bebas [16]. Keduanya tetap melibatkan satu variabel terikat. Persamaan regresi linear adalah sebagai berikut:

$$y = a + b(x) \quad (1)$$

Teknik pemisahan datasets yang digunakan pada penelitian ini adalah *K-Fold Cross Validation*. *Cross Validation* (CV) merupakan teknik yang digunakan untuk melakukan evaluasi terhadap kemampuan suatu model persamaan (statistik) atau algoritma yang dihasilkan dari suatu dataset yang dipisahkan menjadi dua subsets, yaitu training data dan testing data [17]. Training data digunakan untuk melatih model yang dihasilkan untuk mampu mengenali pola data, sedangkan testing data digunakan untuk melakukan percobaan terhadap model yang dihasilkan tersebut.

Pada umumnya, *Cross Validation* (CV) digunakan untuk mendeteksi overfitting atau underfitting pada model persamaan saat pemisahan dataset dan mengetahui keakuratan model tersebut [18]. Salah satu teknik *Cross Validation* (CV) yang banyak digunakan adalah *K-Fold* sehingga disebut *K-Fold Cross Validation*.

Dataset dipisahkan secara terstruktur menjadi *k-folds*. Nilai *k* merujuk pada besar pemisahan dataset menjadi training data dan testing data. Pemisahan dataset ini dilakukan dengan membagi dataset menjadi *k* bagian

yang sekiranya sama rata. Nilai k yang dapat digunakan adalah 1, 2, 3, ... n , tetapi nilai k yang biasanya digunakan adalah 5 atau 10 [19].

Metode pengujian model yang dilakukan adalah uji penyimpangan data. Hasil prediksi dikatakan tidak harus memberikan jawaban yang pasti, tetapi diharapkan mampu memberikan jawaban yang sedekat mungkin. Hal ini dapat diukur dengan mendapatkan nilai ketepatan hasil prediksi [20]. Ketepatan hasil prediksi ini dapat dilakukan dengan menguji penyimpangan data yang akan menghasilkan seberapa besar kesalahan (error) dari dataset yang digunakan. Pengukuran atau perhitungan untuk uji penyimpangan data dapat dilakukan dengan dua model evaluasi, yaitu *Root Mean Squared Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE).

Root Mean Squared Error (RMSE) merupakan salah satu metode evaluasi model dari regresi linear dengan mengukur keakuratan hasil prediksi yang dilakukan dari perhitungan nilai MSE [21]. Semakin rendah nilai RMSE, semakin baik prediksi yang dihasilkan. Nilai rendah yang dimaksud adalah nilai RMSE yang mendekati nol (0) yang mana menunjukkan bahwa hasil prediksi yang dilakukan semakin akurat [22].

Mean Absolute Percentage Error (MAPE) juga merupakan salah satu evaluasi model dari regresi linear untuk mengukur besar kesalahan rata-rata mutlak dalam persentase [23]. MAPE menjadi lebih mudah dipahami karena hasil perhitungan berupa nilai persentase besarnya kesalahan (*error*) data. Prinsip pengujian MAPE ini sama dengan RMSE yang adalah semakin kecil nilai persentase kesalahan, semakin akurat hasil prediksi yang dilakukan. Rentang nilai persentase MAPE dijabarkan sebagai berikut:

1. Nilai MAPE $<10\%$, maka kemampuan model prediksi sangat baik;
2. Nilai MAPE $10\% - 20\%$, maka kemampuan model prediksi baik;

3. Nilai MAPE $20\% - 50\%$, maka kemampuan model prediksi layak; dan
4. Nilai MAPE $>50\%$, maka kemampuan model prediksi buruk.

D. Instrumen Penelitian

Instrumen yang digunakan dalam penelitian ini adalah bahasa pemrograman Python melalui *Google Collaboration*. Python adalah bahasa yang ditemukan oleh seorang pengembang perangkat lunak, Guido van Rossum, pada tahun 1991. Tujuan awalnya adalah untuk membuat segala sesuatu menjadi lebih mudah bagi pengembang (developers). Oleh karena itu, Python dapat digunakan untuk membangun suatu aplikasi, seperti website atau aplikasi perusahaan [24]. Python merupakan bahasa pemrograman tingkat tinggi (high-level) dengan sistem penafsiran (interpreted) untuk berbagai tujuan (general-purpose). Digunakan untuk berbagai tujuan berarti Python fleksibel yang mana memiliki operasi yang dapat diaplikasikan secara luas sehingga mudah beradaptasi dengan suatu domain [25]. Dalam bidang *Data Science* dan *Analysis*, Python banyak digunakan untuk membuat suatu model algoritma machine learning, melakukan manipulasi dan analisis data, melakukan visualisasi data, serta melakukan kalkulasi statistik yang sifatnya kompleks [26].

Algoritma yang digunakan adalah Algoritma Regresi Linear Sederhana untuk menghasilkan model persamaan bagi prediksi. Teknik pemisahan data menjadi training data dan testing data adalah *K-Fold Cross Validation* dengan nilai k yang digunakan adalah 3, 5, 7, dan 9 dengan masing-masing iterasi foldnya.

E Tahapan Penelitian

Tahapan penelitian merupakan langkah-langkah yang dilakukan selama penelitian berlangsung untuk mencapai tujuan penelitian. Adapun tahapan yang dilakukan pada penelitian ini sebagai berikut: (1) Mengidentifikasi masalah. (2) Melakukan studi literatur. (3) Melakukan perumusan masalah. (4) Melakukan

pengumpulan data dari dua tahun terakhir yaitu 2021 – 2022. (5) Melakukan *preprocessing* data. (6) Melakukan analisis data dengan pendekatan *Exploratory Data Analysis* (EDA). (7) Melakukan prediksi menggunakan model persamaan dengan memasukkan variabel x untuk bulan Juni sampai dengan Desember 2023; (8) Melakukan uji penyimpangan data dengan RMSE dan MAPE; dan (9) Menarik kesimpulan dari penelitian.

III. HASIL DAN PEMBAHASAN

A. Hasil Prediksi Per Produk

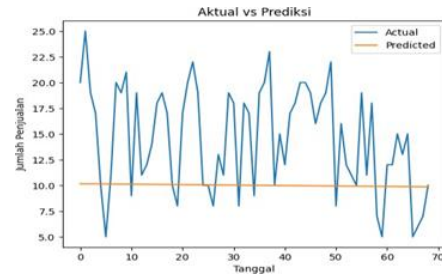
Seluruh datasets beberapa produk makanan dan minuman yang sudah melalui tahap *preprocessing* data, kemudian dianalisis untuk mendapatkan suatu model persamaan yang akan menghasilkan prediksi jumlah penjualan untuk bulan Juni sampai dengan Desember 2023. Tahapan analisis data yang dilakukan adalah percobaan untuk masing-masing nilai k, yaitu 3, 5, 7, dan 9 beserta masing-masing iterasinya dengan jumlah yang sama. Adapun hasil prediksi dari beberapa produk (sebagai perwakilan) sebagai berikut:

1. Hasil Prediksi Seblak

Dari hasil analisis, model persamaan regresi linear yang digunakan berasal dari Percobaan 4 dengan k = 9 dan Fold = 2. Model persamaan yang dihasilkan adalah:

$$y = 10.464911504576325 + (-0.00438354) (x).$$

Konstanta a bernilai positif dan b bernilai negatif sehingga hubungan keduanya berlawanan yang mana mengakibatkan prediksi jumlah penjualan seblak cenderung menurun. Adapun visualisasi perbandingan data aktual dengan data hasil percobaan prediksi (testing data) seblak sebagai berikut.



Gambar 1 Visualisasi Prediksi Jumlah Penjualan Seblak

Dengan demikian, hasil prediksi jumlah penjualan seblak di bulan Juni sampai dengan Desember adalah sebagai berikut.

TANGGAL	JUMLAH PENJUALAN	
0	771	7.085202
1	772	7.080818
2	773	7.076435
3	774	7.072051
4	775	7.067668
...
209	980	6.169042
210	981	6.164658
211	982	6.160275
212	983	6.155891
213	984	6.151508

214 rows x 2 columns

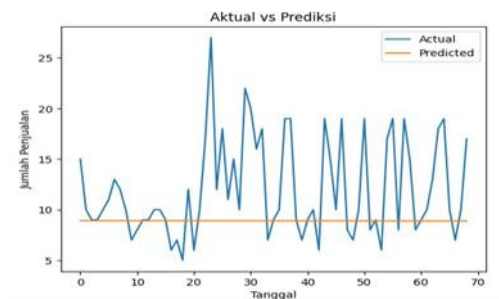
Gambar 2 Hasil Prediksi Jumlah Penjualan Seblak

Berdasarkan hasil prediksi jumlah penjualan seblak tersebut, jumlah penjualan seblak di bulan Juni sampai dengan Desember 2023 mengalami penurunan meskipun tidak signifikan.

2. Hasil Prediksi Dimsum

Dari hasil analisis, model persamaan regresi linear yang digunakan berasal dari Percobaan 4 dengan k = 9 dan Fold = 1. Model persamaan yang dihasilkan adalah:

$$y = 8.911359718995316 + (-0.00054492) (x).$$



Gambar 3 Visualisasi Prediksi Jumlah Penjualan Dimsum

Konstanta a bernilai positif dan b bernilai negatif sehingga hubungan keduanya berlawanan yang mana mengakibatkan prediksi jumlah penjualan dimsum cenderung menurun. Adapun visualisasi perbandingan data aktual dengan data hasil percobaan prediksi (testing data) dimsum sebagai berikut.

Dengan demikian, hasil prediksi jumlah penjualan dimsum di bulan Juni sampai dengan Desember adalah sebagai berikut.

TANGGAL	JUMLAH PENJUALAN
0	771
1	772
2	773
3	774
4	775
...	...
209	980
210	981
211	982
212	983
213	984

214 rows x 2 columns

Gambar 4 Hasil Prediksi Jumlah Penjualan Dimsum

Berdasarkan hasil prediksi jumlah penjualan dimsum tersebut, jumlah penjualan dimsum di bulan Juni sampai dengan Desember 2023 mengalami penurunan meskipun tidak signifikan.

B. Uji Root Mean Squared Error (RMSE)

Setelah analisis data dan prediksi jumlah penjualan dilakukan, hasil prediksi dari penggunaan model persamaan tersebut dievaluasi menggunakan pengujian penyimpangan data, salah satunya adalah RMSE. Adapun hasil perhitungan RMSE menggunakan fungsi Python dari setiap dataset produk sebagai berikut.

1. Hasil RMSE Seblak

Tabel 1 Hasil RMSE Dataset Penjualan Seblak

Folds	k = 3	k = 5	k = 7	k = 9
1	6.98	9.33	7.49	6.78
2	5.22	7.00	5.78	6.76
3	4.13	4.17	5.95	6.75
4	-	5.73	3.96	4.99
5	-	2.97	7.08	4.06
6	-	-	3.32	6.76
7	-	-	2.90	4.86

8	-	-	-	2.84
9	-	-	-	3.01

Berdasarkan Tabel 1, nilai RMSE untuk prediksi penjualan seblak dengan Percobaan 4 untuk k = 9 dan Fold = 2 adalah 6,76.

2. Hasil RMSE Dimsum

Tabel 2 Hasil RMSE Dataset Penjualan Dimsum

Folds	k = 3	k = 5	k = 7	k = 9
1	5.92	8.03	6.73	5.71
2	5.22	6.39	5.30	6.39
3	3.27	4.61	5.37	6.26
4	-	5.19	4.13	4.85
5	-	2.83	6.79	3.71
6	-	-	3.37	7.08
7	-	-	2.66	4.17
8	-	-	-	2.67
9	-	-	-	2.78

Berdasarkan Tabel 2, nilai RMSE untuk prediksi penjualan dimsum dengan Percobaan 4 untuk k = 9 dan Fold = 1 adalah 5.71.

C. Uji Mean Absolute Percentage Error (MAPE)

Selain RMSE, uji penyimpangan data juga dapat dilakukan menggunakan MAPE. Hasil perhitungan MAPE ini berupa persentase kesalahan data (penyimpangan data) sehingga lebih mudah dipahami pembacaanya. Adapun hasil perhitungan MAPE menggunakan fungsi Python dari setiap dataset produk sebagai berikut.

1. Hasil MAPE Seblak

Tabel 3 Hasil MAPE Dataset Penjualan Seblak

Folds	k = 3	k = 5	k = 7	k = 9
1	71.02%	49.56%	41.08%	40.98%
2	106.84%	206.79%	94.00%	35.86%
3	59.97%	59.98%	180.35%	212.97%
4	-	92.74%	68.77%	128.93%
5	-	65.13%	100.38%	68.14%
6	-	-	76.31%	117.32%
7	-	-	57.66%	66.27%

8	-	-	-	62.87%
9	-	-	-	66.57%

Berdasarkan Tabel 3, dari empat percobaan yang dilakukan dengan nilai k yang berbeda dan masing-masing folds, nilai MAPE paling kecil terdapat pada Percobaan 4.2 dengan nilai k = 9 dan fold = 2, yaitu 35,86%.

Nilai MAPE tersebut berada pada range MAPE 20% – 50%. Artinya, kemampuan model persamaan dari nilai k = 9 dan Fold.

2. Hasil MAPE Dimsum

Tabel 4 Hasil MAPE Dataset Penjualan Dimsum

Folds	k = 3	k = 5	k = 7	k = 9
1	67.42%	43.35%	32.76%	28.17%
2	83.19%	185.54%	97.70%	36.09%
3	55.83%	48.87%	148.33%	200.24%
4	-	81.31%	54.34%	126.65%
5	-	51.59%	81.38%	31.70%
6	-	-	72.93%	95.70%
7	-	-	43.06%	72.69%
8	-	-	-	48.47%
9	-	-	-	46.52%

Berdasarkan Tabel 4, dari empat percobaan yang dilakukan dengan nilai k yang berbeda dan masing-masing folds, nilai MAPE paling kecil terdapat pada Percobaan 4.1 dengan nilai k = 9 dan Fold = 1, yaitu 28,17%.

Nilai MAPE tersebut berada pada range MAPE 20% – 50%. Artinya, kemampuan model persamaan dari nilai k = 9 dan Fold = 1 layak digunakan untuk prediksi jumlah penjualan dimsum pada bulan Juni sampai dengan Desember 2023.

IV. SIMPULAN

Berdasarkan hasil penelitian, dapat ditarik kesimpulan sebagai berikut. Kantin Berkat merupakan UMKM di bidang kuliner yang menjual makanan dan minuman rumahan. Selama beroperasi, Kantin Berkat mengalami masalah yang mana data penjualan produknya fluktuatif sehingga jumlah penjualan selanjutnya tidak bisa ditentukan secara gamblang yang berakibat banyak produk tidak laku dan harus dibuang.

Oleh karena itu, dilakukan prediksi jumlah penjualan untuk beberapa produk Kantin Berkat menggunakan Algoritma Regresi Linear dengan pengkodean pada Python melalui *Google Collaboration* dan menggunakan teknik pemisahan data, yaitu *K-Fold Cross Validation*. Data penjualan yang digunakan merupakan data historis penjualan beberapa produk selama dua tahun terakhir (2021 – 2022), yaitu seblak, dimsum, singkong, es jeruk, dan es teh manis.

Berdasarkan hasil analisis data dan hasil perhitungan prediksi, dapat ditarik kesimpulan sebagai berikut.

- Kelima produk yang dilakukan prediksi menggunakan model persamaan dengan nilai k = 9. Artinya, seluruh produk menggunakan model persamaan pada Percobaan 4 berdasarkan pada nilai MAPE paling kecil.
- Berdasarkan hasil perhitungan prediksi, jumlah penjualan seblak, dimsum, singkong, dan es jeruk pada bulan Juni sampai dengan Desember 2023 mengalami penurunan. Penurunan paling signifikan terjadi pada produk singkong dengan jumlah penjualan bernilai negatif. Dengan demikian, keempat produk tersebut akan memberikan kerugian bagi Kantin Berkat. Di sisi lain, jumlah penjualan es teh manis pada bulan Juni sampai dengan Desember 2023 diprediksi mengalami kenaikan meskipun tidak signifikan. Dengan demikian, penjualan es teh manis ini akan memberikan sedikit keuntungan bagi Kantin Berkat.
- Nilai penyimpangan data, MAPE, produk seblak, dimsum, dan singkong berada pada range 20% – 50% yang artinya model persamaan yang digunakan untuk prediksi adalah layak. Untuk produk es jeruk dan es teh manis memiliki nilai MAPE pada range 10% – 20% yang artinya model persamaan yang

digunakan untuk prediksi adalah baik.

DAFTAR RUJUKAN

- [1] F. Rohmawati, G. Rohman, And S. Mujilahwati, “Sistem Prediksi Jumlah Pengunjung Wisata Wego Kec.Sugio Kab.Lamongan Menggunakan Metode Fuzzy Time Series,” *Jouticla*, Vol. 3, No. 2, 2017.
- [2] Ervina, “Klasifikasi Data: Pengertian, Jenis, Hingga Metodenya,” *Insight Talenta*, Nov. 27, 2019. *Klasifikasi Data: Pengertian, Jenis, Hingga Metodenya* (accessed Feb. 02, 2023).
- [3] T. Löfgren, D. Gravem, and G. Haraldsen, “A glimpse into the businesses’ use of internal and external data sources in decision-making processes,” 2011. [Online]. Available: <http://www.blue-ets.istat.it/index.php?id=39>
- [4] Y. Arisandy And R. Satriawan, “Promosi Dalam Meningkatkan Volume Penjualan Tinjauan Manajemen Syariah,” *Al-Intaj*, Vol. 4, No. 1, Pp. 74–88, 2018.
- [5] “Pengertian UMKM Menurut Undang- Undang, Kriteria, dan Ciri-Ciri UMKM.” <https://sukorejo.semarangkota.go.id/umkm> m (accessed Feb. 02, 2023).
- [6] Yazfinedi, “Usaha Mikro, Kecil, Dan Menengah Di Indonesia: Permasalahan Dan Solusinya,” *Jurnal Ilmiah Kesejahteraan Sosial*, Vol. 14, No. 25, Jun. 2018.
- [7] S. Al Farisi And M. Iqbal Fasa, “Peran Umkm (Usaha Mikro Kecil Menengah) dalam Peningkatkan Kesejahteraan Masyarakat,” *Jurnal Dinamika Ekonomi Syariah*, Vol. 9, No. 1, Pp. 73–84, 2022, [Online]. Available: [Http://ejournal.laipd-nganjuk.ac.id/index.php/es/index](http://ejournal.laipd-nganjuk.ac.id/index.php/es/index)
- [8] Nurmala et al., “Usaha Kuliner Sebagai Penggerak UMKM Pada Masa Pandemi Covid 19,” *STEBIS IGM*, vol. 3, no. 1, pp. 65–74, Jul. 2022.
- [9] E. Prasetyowati, “Aplikasi Penentuan Harga Pokok Produksi Batik Madura Dengan Metode Activity Based Costing Dan Analisis Regresi LINIER,” *JUTI: Jurnal Ilmiah Teknologi Informasi*, vol. 16, no. 1, pp. 48–58, 2018.
- [10] W. Suhermawan, “Analisa Prediksi Penjualan I-Voucher Pada Pt. Indomarcoprismatama Menggunakan Algoritma Regresi Linear Dan Neural Network.”
- [11] I. A. Prakoso, Kusnadi, And B. Nugraha, “Peramalan Penjualan Produk Dengan Metode Regresi Linear Dan Aplikasi POM-QM DI PT XYZ,” *Jurnal Ilmiah Widya Teknik*, vol. 20, no. 1, pp. 17–20, 2021.
- [12] M. Kafil, “Penerapan Metode K- Nearest Neighbors Untuk Prediksi Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso,” *Jurnal Mahasiswa Teknik Informatika*, Vol. 3, No. 2, Pp. 59–66, 2019.
- [13] M. Wijayanti, “Analisa Peramalan Penjualan Produk Susu Kental Manis Carnation Pada Cv Pangan Makmur Irja Sorong.”
- [14] D. Sofita, D. Yuniarti, and R. Goejantoro, “Analisis Regresi Eksponensial (Studi Kasus: Data Jumlah Penduduk dan Kelahiran di Kalimantan Timur pada Tahun 1992-2013),” *Jurnal EKSPONENSIAL*, vol. 6, no. 1, pp. 57–64. May 2015.
- [15] O. Januari Ababil, S. Adi Wibowo, And H. Zulfia Zahro, “Penerapan Metode Regresi Linier Dalam Prediksi Penjualan Liquid Vape Di Toko Vapor Pandaan Berbasis Website,” *Jurnal Mahasiswa Teknik Informatika*, Vol. 6, No. 1, Pp. 186–195, Feb. 2022.
- [16] S. Disa, “Penerapan Metode Regresi Linear Dalam Pembuatan Perangkat Lunak Simulasi Target Penjualan,” Pp. 82–89.
- [17] L. Octoria Sitanggang and N. Bahtiar, “Aplikasi Data Mining Untuk Mendeteksi Penyakit Hepatitis Menggunakan Metode Support Vector Machine (SVM) Polynomial (Studi Kasus : Data Pasien Hati India),” *Jurnal Masyarakat Informatika*, vol. 10, no. 1, pp. 20–27.
- [18] H. Azis, P. Purnawansyah, F. Fattah, and I. P. Putri, “Performa Klasifikasi K-NN dan *Cross Validation* Pada Data Pasien Pengidap Penyakit Jantung,” *ILKOM Jurnal Ilmiah*, vol. 12, no. 2, pp. 81–86, Aug. 2020, doi: 10.33096/ilkom.v12i2.507.81-86.
- [19] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning Data Mining, Inference, and Prediction*, Second. Springer.
- [20] K. Margi And S. Pendawa, “Analisa Dan Penerapan Metode Single Exponential Smoothing Untuk Prediksi Penjualan Pada Periode Tertentu (Studi Kasus : Pt. Media Cemara Kreasi),” In *Prosiding Snatif*, 2015, Pp. 259–266.
- [21] F. Widiastuti, W. Murniati, And Saikin, “Penerapan Data Mining Untuk Memprediksi Penjualan Kain Tenun Mnggunakan Regresi Linear Studi Kasus: Ud.Bintang Remawe Sukarare,” *Jurnal Ilmiah Teknik Mesin, Elektro, Dan Komputer*, Vol. 2, No. 1, Pp. 27–39, 2022.
- [22] L. Y. Kurniawati, H. Tjandrasa, And I. Ariesanti, “Prediksi Pergerakan Harga Saham Menggunakan Support Vector Regression,” *Jurnal Simantec*, Vol. 4, No. 1, Pp. 35–47, Jun. 2014.
- [23] I. Nabillah and I. Ranggadara, “*Mean Absolute Percentage Error* untuk Evaluasi Hasil Prediksi Komoditas Laut,” *JOINS (Journal of Information*

- System), vol. 5, no. 2, pp. 250–255, Nov. 2020, doi: 10.33633/joins.v5i2.3900.
- [24] Steve Safarowic, “Why is Python Considered a High-Level Programming Language?,” 2022. <https://xccelerate.co/blog/why-python-a-high-level-programming-language> (accessed Jul. 18, 2023).
- [25] J. V. Guttag, Introduction to Computation and Programming Using Python with Application to Computational Modeling and Understanding Data, vol. 3. 2021.
- [26] T. Kurnialensya, “Kegunaan & Fungsi Pemrograman Python Analisis Data dan Machine Learning,” Jul. 15, 2022. <https://teknik-informatika.s1.stekom.ac.id/informasi/baca/Kegunaan-Fungsi-Pemrograman-Python-Analisis-Data-dan-Ma>