

## SISTEM PREDIKSI PENJUALAN TRIPLEK MENGGUNAKAN METODE REGRESI TIME SERIES K-NEAREST NEIGHBOR (KNN) PADA TOKO MAKMUR CIBUBUR

Michael Joses Dharmali<sup>1</sup>, Desi Arisandi<sup>2</sup>, Tri Sutrisno<sup>3</sup>

Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara  
Jln. Letjen S. Parman No. 1, Jakarta, 11440, Indonesia

E-mail: [michaeljosesd@gmail.com](mailto:michaeljosesd@gmail.com)<sup>1</sup>, [desia@fti.untar.ac.id](mailto:desia@fti.untar.ac.id)<sup>2</sup>, [tris@fti.untar.ac.id](mailto:tris@fti.untar.ac.id)<sup>3</sup>

### ABSTRACT

*The availability of goods holds a very important role in the trading industry. This also applies to the availability of plywood at Toko Makmur Cibubur. The purpose of this research is to create a program that is able to predict plywood sales using the K-Nearest Neighbor (KNN) regression algorithm on time series data. The research method is quantitative by using time series data on sales of seven types of plywood. The seven models that have been produced in this study achieve the highest average accuracy rate of 62.71% for 28 validation data. The results and conclusions that can be drawn from this study are that the KNN method can predict plywood sales in a time series with sufficient accuracy, so that it can help determine plywood supplies.*

**Keywords:** *k-nearest neighbor, regression, time series, prediction, sales, plywood*

### ABSTRAK

Memastikan ketersediaan barang memiliki peranan yang sangat penting dalam perusahaan dagang. Hal ini juga berlaku pada ketersediaan triplek dalam Toko Makmur Cibubur. Tujuan dari penelitian ini adalah pembuatan program yang dapat memprediksi penjualan triplek menggunakan algoritma regresi *K-Nearest Neighbor* (KNN) pada data *time series*. Metode penelitian bersifat kuantitatif dengan menggunakan data *time series* penjualan tujuh jenis triplek. Ketujuh model yang telah dihasilkan dalam penelitian ini mencapai rata-rata tingkat akurasi tertinggi sebesar 62,71% untuk 28 data validasi. Hasil dan kesimpulan yang dapat diambil dari penelitian ini adalah bahwa metode KNN dapat memprediksi penjualan triplek bersifat *time series* dengan akurasi yang cukup, sehingga dapat membantu menentukan persediaan triplek.

**Kata kunci:** *K-Nearest Neighbor, Regresi, Time Series, Prediksi, Penjualan, Triplek*

### 1. PENDAHULUAN

Kemajuan teknologi informasi yang telah dicapai hingga saat ini telah membawa evolusi pesat pada setiap aspek kehidupan manusia, mulai dari aspek pendidikan hingga ke aspek bisnis dan perusahaan. Semakin berkembangnya bisnis dan teknologi informasi, semakin banyak perubahan terhadap kuantitas dan variasi yang ada dalam kebutuhan bisnis.

Agar dapat memenuhi kebutuhan bisnis dengan cepat dan baik, dibutuhkan pengolahan data dan informasi dengan menggunakan sistem komputer (Kurniawan, 2022). Pada zaman yang sedang berkembang pesat ini, dibutuhkan adanya sistem komputer dan otomatisasi bisnis agar bisa memperoleh untung dari proses transaksi untuk terus mengembangkan dan memajukan bisnis.

Toko Makmur merupakan toko yang telah bergerak dalam penjualan berbagai macam jenis triplek dan blockboard selama 2 tahun. Toko ini berlokasi di Ciangsana, Gunung Putri, Bogor, Jawa Barat. Toko ini dapat melayani permintaan distribusi untuk pertokoan dan eceran untuk perorangan. Adapun sebagian besar jenis triplek yang dijual

memiliki variasi berdasarkan ketebalan, ukuran, dan jenis bahan dasar kayu yang digunakan.

Proses transaksi dan inventori Toko Makmur yang masih sepenuhnya manual menyebabkan sering terjadinya kekeliruan dalam pencatatan stok barang yang dimiliki, sehingga sistem penjualan dan penyetokan triplek tidak dapat berjalan dengan optimal. Sistem penyetokan yang mengikuti acuan harga pasaran dan tingkat permintaan toko tersebut menyebabkan toko kehilangan potensi pelanggan, potensi pemasukkan dan laba, serta ketidakpuasan pelanggan. Ketika permintaan pasar terhadap suatu jenis triplek sedang meningkat dan toko sedang kekurangan / kehabisan stok untuk jenis tersebut, pelanggan seringkali merasa tidak puas dan pergi untuk mencari toko triplek lain.

Tidak jarang terjadi, salah satu jenis triplek dapat mengalami kondisi pasar yang sedang lesu. Hal tersebut dapat mengakibatkan turunnya angka permintaan oleh pelanggan, dan menimbulkan penumpukan jumlah triplek toko dalam durasi waktu yang lama. Tumpukan triplek yang tidak dirawat dapat mengalami penurunan kualitas, hingga kerusakan total. Sehingga, nilai jual triplek tersebut

dapat mengalami penurunan atau bahkan tidak dapat dijual lagi (Qossam, 2019).

Dalam peramalan penjualan triplek, metode K-Nearest Neighbor (KNN) dapat digunakan. Metode tersebut selain dapat digunakan untuk membuat suatu klasifikasi, dapat juga digunakan untuk membuat peramalan dengan regresi. Metode KNN dalam peramalan data *time series* triplek memiliki keunggulan dalam menangani data berdimensi tinggi dan data yang tidak penuh / *incomplete*, seperti data penjualan triplek (Resti, 2022).

Oleh sebab itu, penelitian ini bertujuan untuk merancang aplikasi untuk membuat prediksi / peramalan jumlah penjualan triplek yang akan terjual untuk beberapa waktu ke depan dengan menggunakan metode KNN. Diharapkan dengan prediksi penjualan ini, pemilik Toko Makmur dapat mengoptimalkan penjualan dan kualitas triplek yang dimiliki.

## 2. METODE

### 2.1 Data

Data yang digunakan dalam penelitian ini berasal dari observasi langsung melalui data penjualan tujuh jenis triplek yang terjadi di Toko Makmur yang terjadi selama 2 tahun, yaitu tahun 2020 hingga tahun 2022. Data tersebut berjumlah 140 baris dengan 8 kolom fitur yang meliputi tanggal dan jumlah penjualan tiap triplek per minggu. Jenis triplek yang digunakan dalam pembuatan model perancangan sistem ini menggunakan jenis yang memiliki frekuensi penjualan yang cukup dan tinggi. Jenis triplek tersebut adalah sebagai berikut:

**Tabel 1.** Jenis triplek

No.	Jenis Triplek
1.	Triplek 3mm 4x8 Alba
2.	Triplek 3mm 4x8 Single Face Melamin
3.	Triplek 9mm 4x8 Alba
4.	Triplek 9mm 4x8 Meranti Campur
5.	Triplek 12mm 4x8 Alba
6.	Triplek 15mm 4x8 Alba
7.	Triplek 15mm 4x8 Meranti Campur

Data riwayat penjualan setiap jenis triplek memiliki struktur data berkala dengan hanya satu fitur, sehingga diperlukan pengolahan tambahan agar data berkala dapat memiliki pasangan fitur / input dan hasil / output. Hal ini dilakukan agar data dapat terbaca dan diproses oleh Python.

Pengolahan pasangan fitur pada data dilakukan dengan menggunakan fitur *time series generator* yang terdapat pada pustaka Keras. Data riwayat penjualan triplek akan memiliki lag sepanjang *time step* t sebagai fitur, dan nilai time step pada waktu sekarang sebagai hasil.

### 2.2 Algoritma

Untuk meramal penjualan triplek, pada penelitian ini digunakan metode regresi k-Nearest Neighbor (KNN). KNN merupakan metode pembelajaran mesin yang umum digunakan untuk pengenalan pola / *pattern recognition* dan klasifikasi objek (Stephensen, 2022).

Kalkulasi dalam klasifikasi KNN bekerja berdasarkan jarak data pelatihan / training terdekat. Untuk setiap data baru yang belum diketahui sebelumnya, KNN akan melakukan prediksi klasifikasi data tersebut berdasarkan klasifikasi data lain yang memiliki jarak terdekat dengannya. Selanjutnya, dilakukan penghitungan jarak antara data baru yang belum diketahui dengan semua data yang telah diketahui dalam data pelatihan. Data / tetangga dengan jarak terkecil menjadi penentu hasil klasifikasi data baru tersebut.

**Gambar 1.** Ilustrasi jarak dalam KNN

Pada umumnya, penghitungan jarak dalam KNN menggunakan metode Jarak Euclidean atau sering disebut dengan Euclidean distance (Natcha et al., 2021). Jarak Euclidean berfungsi untuk memberi gambaran jarak antara dua objek dalam satuan angka / numerik. Adapun rumus Jarak Euclidean adalah sebagai berikut:

$$D(X_1, X_2) = \sqrt{\sum_{i=1}^n (X_{1i} - X_{2i})^2} \quad (1)$$

Dengan keterangan:

D = distance / jarak

$X_{1i}$  = poin data training / pelatihan ke-i

$X_{2i}$  = poin data testing / pengujian ke-i

n = jumlah data yang ada

Hasil D yang semakin besar menandakan jarak antara data pelatihan dan data pengujian yang jauh. Sebaliknya, hasil D yang semakin kecil menandakan bahwa jarak semakin dekat. Penentuan nilai k / banyaknya data tetangga yang mendasari jauh / dekatnya jarak tersebut merupakan proses uji coba (*trial and error*) dan bergantung pada besarnya data yang diproses.

Selain untuk kasus klasifikasi, KNN juga dapat digunakan dalam kasus regresi / prediktif. Proses dalam perhitungan regresi menggunakan langkah yang sama dengan kasus klasifikasi, tetapi hasil

akhir dari regresi akan berupa rata-rata dari tetangga k terdekat (Abi et al., 2019). Rumus regresi dari KNN adalah sebagai berikut di bawah:

$$\hat{y}(X) = \frac{1}{K} \sum_{i=1}^k Y_i \quad (2)$$

Dengan keterangan:

y = Hasil prediksi

K = nilai k / tetangga dalam KNN

$Y_i$  = data pengujian ke - i.

Langkah – langkah untuk menghitung nilai prediksi *time series* dengan menggunakan regresi KNN ke dalam beberapa poin berikut di bawah (Marinoiu, 2018):

1. Membaca data dan menentukan nilai p dan k.
2. Buat transformasi data ke dalam bentuk tabel dengan jumlah fitur sejumlah p.
3. Menghitung jarak Euclidean antara data pengujian (data yang akan diprediksi) dengan data pelatihan (data yang ada sebelumnya).
4. Mengurutkan hasil jarak Euclidean dari jarak terkecil ke data terbesar.
5. Menghitung rata – rata target dari k data dengan jarak terdekat.

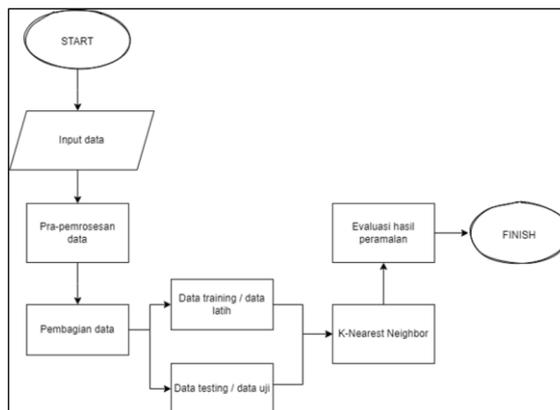
Untuk data *time series* dengan satu fitur, diperlukan rekonstruksi kembali pada data ke dalam bentuk yang dapat diterima dan diolah menggunakan suatu algoritma pembelajaran yang diawasi atau sering disebut dengan supervised learning. Data *time series* perlu direpresentasikan ke dalam sebuah tabel dengan kolom – kolom yang merepresentasikan masukkan / input fitur dan keluaran / output yang merupakan suatu target (Seruni et al., 2020).

Dalam artikel yang sama, Cristian menggambarkan transformasi data tersebut dengan anggapan bahwa data *time series* terdiri atas suatu tuple  $y = \{y_1, y_2, y_3, y_4, \dots, y_n\}$ . Tuple tersebut kemudian akan dilakukan transformasi ke dalam bentuk sebuah tabel berisi kolom input fitur sejumlah p dan kolom output target y.

**Tabel 2.** Ilustrasi rekonstruksi data *time series* menjadi *lag*

P1	P2	P3	Target
y1	y2	y3	y4
y5	y6	y7	y8
y9	y10	y11	y12
y13	...	...	yn

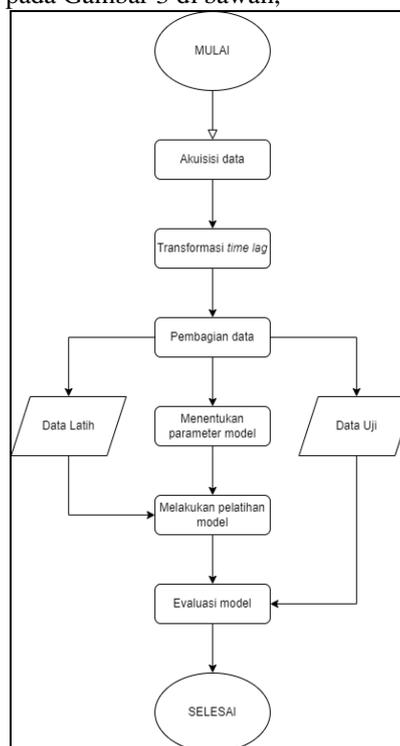
Berikut merupakan diagram alir metode KNN pada Gambar 2 di bawah untuk menjelaskan lebih lanjut bagaimana metode tersebut dapat membuat prediksi dari data.



**Gambar 2.** Diagram alir metode KNN

### 2.3 Rancangan Eksperimen

Untuk Berikut merupakan diagram alir yang menggambarkan proses berjalannya penelitian seperti pada Gambar 3 di bawah;



**Gambar 3.** Diagram alir rancangan penelitian

Berdasarkan gambar tersebut, tahap pertama yang harus dilakukan adalah dengan mengakuisisi data penjualan ketujuh jenis triplek dari observasi langsung terhadap buku transaksi Toko Makmur Cibubur. Setelah data observasi tersebut diketik secara manual melalui perangkat lunak Microsoft Excel ke dalam bentuk lembaran tabel dengan format *xlsx*, maka tahap selanjutnya adalah transformasi data.

Setelah data berhasil dilakukan transformasi ke dalam bentuk *time lag* seperti pada Tabel 2, maka data dapat dibagi ke dalam dataset latih dan uji. Setelah data telah terbagi ke dalam dataset latih dan

uji, maka tahap selanjutnya adalah menentukan parameter model. Penentuan parameter model ini dilakukan dengan proses uji coba (*trial and error*) dengan mencoba satu-satu nilai parameter berbeda terhadap data latih saat pelatihan model. Nilai parameter yang berbeda juga akan menghasilkan prediksi yang berbeda, sehingga hasil evaluasi yang didapat juga akan berbeda.

Setelah parameter model yang optimal telah ditentukan, maka tahap selanjutnya adalah pelatihan model dengan parameter dan dataset latih yang telah dibuat. Model akan memilih nilai parameter dalam proses pengujian yang menghasilkan nilai evaluasi terendah.

Tahap terakhir sebelum model dapat melakukan prediksi adalah proses evaluasi / pengujian. Pengujian melihat perbandingan antara data uji dengan data yang diprediksi untuk melihat akurasi yang dihasilkan model.

## 2.4 Metode evaluasi

Metode yang digunakan untuk melakukan evaluasi pada akurasi model adalah metode *Mean Absolute Percentage Error* (MAPE). MAPE merupakan metode pengukuran batas kesalahan / margin of error dengan mengukur selisih nilai kesalahan absolut yang terjadi antara nilai peramalan dan nilai aktual pada data (Diajeng et al., 2020). Adapun rumus untuk menghitung batas kesalahan dengan MAPE adalah sebagai berikut:

$$MAPE = \sum_{i=1}^n \left( \left| \frac{\hat{y}_i - y_i}{y_i} \right| \right) 100 \quad (3)$$

Dengan keterangan :

- n = Jumlah data dalam periode tertentu
- $\hat{y}_i$  = Nilai peramalan / prediksi yang dihasilkan
- $y_i$  = Nilai aktual yang ada dalam data

Terdapat kriteria dalam nilai MAPE yang diraih dalam perhitungan kesalahan. Kriteria tersebut menentukan seberapa baik / buruknya kemampuan peramalan berdasarkan hasil persentase kesalahan MAPE yang didapat (Sindhoora et al., 2021). Adapun kriteria tersebut terbagi atas beberapa poin yang disajikan ke dalam tabel berikut:

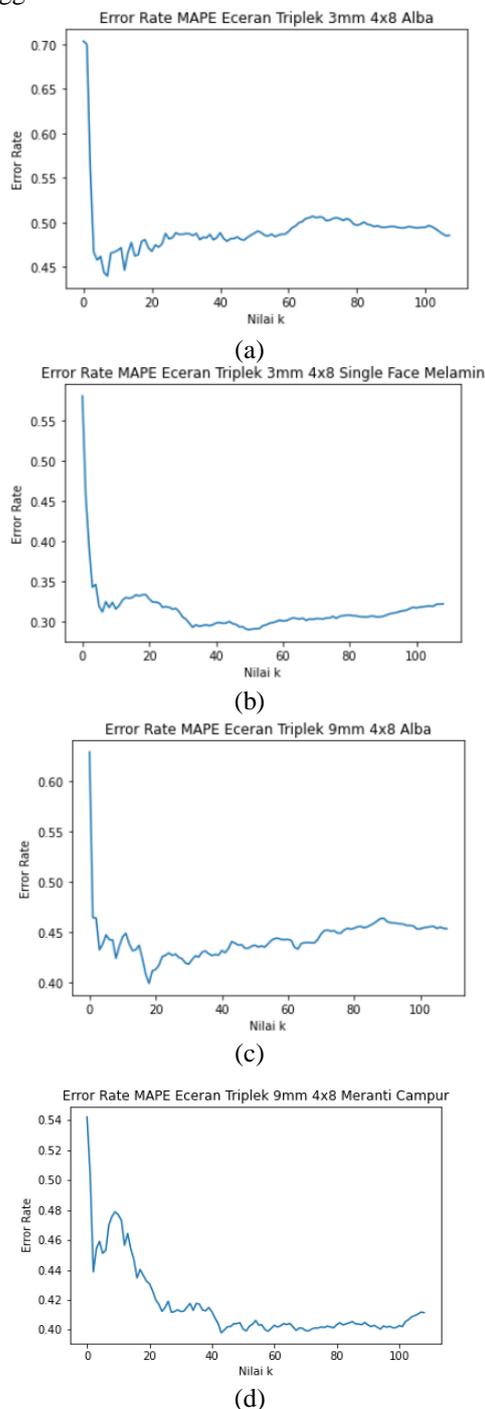
**Tabel 3.** Kriteria hasil MAPE

Nilai MAPE	Kriteria
<10%	Sangat baik / sangat akurat
10% - 20%	Baik
20% - 50%	Cukup
>50%	Kurang / buruk

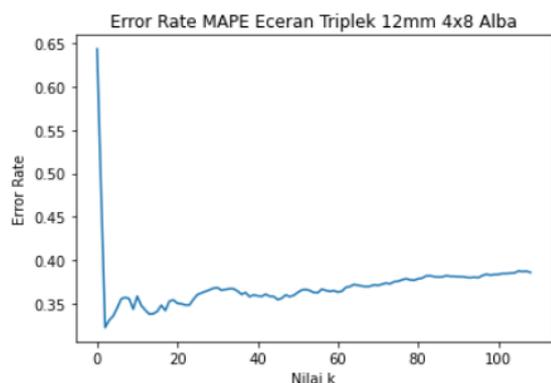
## 3. PEMBAHASAN

Setelah data diubah ke dalam bentuk time lag serta dibagi ke dalam data latih dan uji, terdapat 109 data latih dan 28 data uji untuk setiap jenis triplek.

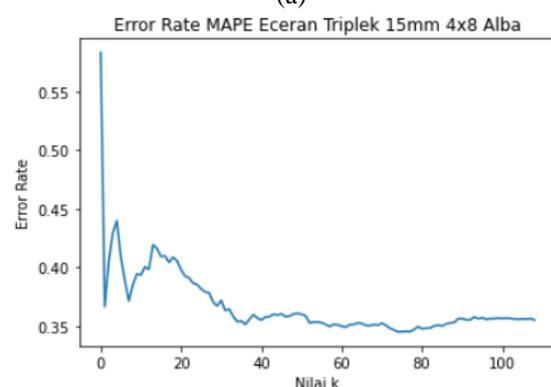
Untuk mencari nilai k yang optimal, dibutuhkan pengujian pengaruh nilai parameter model k yang berbeda-beda terhadap hasil evaluasi MAPE dari ketujuh model yang dibuat. Berikut pada Gambar 4 merupakan hasil plot yang didapat setelah menguji pengaruh parameter k terhadap model triplek 3mm hingga 9mm.



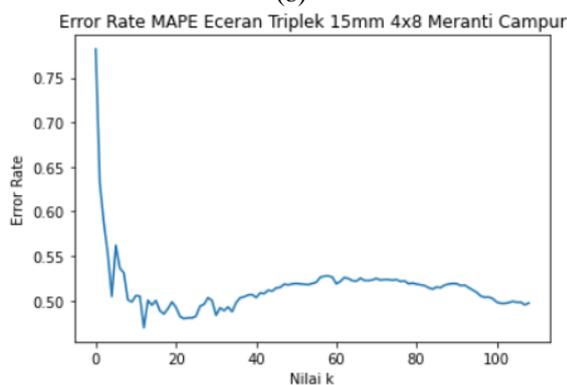
**Gambar 4.** Plot pengaruh nilai k terhadap data triplek 3mm – 9mm (a) Triplek 3mm 4x8 Alba (b) Triplek 3mm 4x8 Single Face Melamin (c) Triplek 9mm 4x8 Alba (d) Triplek 9mm 4x8 Meranti Campur



(a)



(b)



(c)

**Gambar 5.** Plot pengaruh nilai k terhadap data triplek 12mm – 15mm (a) Triplek 12mm 4x8 Alba (b) Triplek 15mm 4x8 Alba (c) Triplek 15mm 4x8 Meranti Campur

Pada Tabel 4 di bawah, telah dijabarkan hasil pengujian berupa nilai k optimal yang menghasilkan MAPE terendah untuk masing-masing model peramalan. Terlihat bahwa rentang nilai k optimal yang dihasilkan tiap model lumayan berbeda jauh satu sama lain. Hal ini dapat disebabkan oleh pola yang berbeda-beda dalam setiap jenis triplek.

**Tabel 4.** Nilai k optimal setiap jenis triplek

Jenis Triplek	Nilai k Optimal	Nilai MAPE
3mm 4x8 Alba	8	44%
3mm 4x8 <i>Single Face</i> Melamin	51	29%
9mm 4x8 Alba	19	40%
9mm 4x8 Meranti Campur	8	35%
12mm 4x8 Alba	3	32%
15mm 4x8 Alba	75	34%
15mm 4x8 Meranti Campur	13	47%

Pada Tabel 5 hingga 8 di bawah, telah dijabarkan sebagian perbedaan jumlah penjualan dari data aktual dengan jumlah penjualan dari hasil prediksi pada triplek 3mm hingga 9mm.

**Tabel 5.** Perbandingan nilai aktual dengan prediksi (Triplek 3mm)

Triplek 3mm 4x8			
Alba		Single Face Melamin	
Aktual	Prediksi	Aktual	Prediksi
15	23	56	26
11	14	7	24
22	23	15	19
20	24	20	32
9	9	21	36
38	13	18	19

**Tabel 6.** Perbandingan nilai aktual dengan prediksi (Triplek 9mm)

Triplek 9mm 4x8			
Alba		Meranti Campur	
Aktual	Prediksi	Aktual	Prediksi
10	11	7	14
23	26	38	11
17	10	20	20
18	22	17	19
18	17	32	23
12	19	28	22

**Tabel 7.** Perbandingan nilai aktual dengan prediksi (Triplek 12mm)

Triplek 12mm 4x8 Alba	
Aktual	Prediksi
16	16
23	22
31	23
27	17
12	26
9	16

**Tabel 8.** Perbandingan nilai aktual dengan prediksi (Triplek 15mm)

Triplek 15mm 4x8			
Alba		Meranti Campur	
Aktual	Prediksi	Aktual	Prediksi
14	11	12	14
10	13	9	7
10	11	12	8
11	5	14	16
15	9	5	6
13	12	12	11

#### 4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat diambil kesimpulan bahwa program prediksi penjualan triplek dengan algoritma regresi KNN pada data *time series* dapat dilakukan dengan rata-rata tingkat akurasi sebesar 62,71% yang cukup untuk ketujuh model yang dibuat. Dengan demikian, metode ini dapat menjadi solusi alternatif dalam membantu pihak Toko Makmur dalam memprediksi penjualan dan menjaga ketersediaan triplek.

Walaupun demikian, kuantitas data yang digunakan dalam penelitian ini masih sedikit dan terbatas. Kualitas data yang digunakan juga masih dapat ditambahkan dengan menambahkan metode tambahan yang tepat untuk mengatasi pola dalam data yang beragam.

Adapun saran dan masukan yang dapat diberikan untuk penelitian selanjutnya yang sejenis adalah untuk menambah kuantitas data dalam dataset, serta penambahan tindakan dan metode yang tepat dalam pra-pemrosesan data *time series*.

#### PUSTAKA

- Kurniawan, F. (2022). Penentuan Persediaan Bahan Baku Kertas pada Perusahaan Manufaktur dengan Metode Long-Short Term Memory. [http://repository.untar.ac.id/27974/4/535170023\\_Febryo%20Kurniawan\\_04%20Abstrak.pdf](http://repository.untar.ac.id/27974/4/535170023_Febryo%20Kurniawan_04%20Abstrak.pdf)
- Qossam, Muhammad, I. A. (2019). Analisis Daya Saing dan Struktur Pasar Kayu Lapis Indonesia di Pasar Internasional. *Jurnal Ilmiah Mahasiswa FEB Universitas Brawijaya*, 8(2).
- Resti, H. & Erna, Z. A. (2022). Implementasi Metode K-Nearest Neighbor untuk Prediksi Penjualan Furniture pada CV Octo Agung Jepara. [http://eprints.dinus.ac.id/18785/2/jurnal\\_17788.pdf](http://eprints.dinus.ac.id/18785/2/jurnal_17788.pdf)
- Stephensen, R. (2022). Perancangan Sistem Rekomendasi dan Reservasi Hotel Menggunakan Metode Case-Based Reasoning (CBR) dengan Algoritma KNN Berbasis Web.

[http://repository.untar.ac.id/27974/4/535170023\\_Stephensen%20Reynard\\_04%20Abstrak.pdf](http://repository.untar.ac.id/27974/4/535170023_Stephensen%20Reynard_04%20Abstrak.pdf)

- Natcha, K. S., Budi, A. D., Tesa, P. N. (2021). Prediksi Barang Keluar TB Wijaya Bangunan Menggunakan Algoritma KNN Regression dengan RStudio. *JISKA*, 4(2).
- Abi, H., Indriani, Fatma & Muliadi (2019). Metode Timeseries K-Nearest Neighbor Regression dalam Prediksi Barang Keluar pada Gudang PT Putra Prenuer Banjarbaru. *Seminar Nasional Ilmu Komputer (SOLITER)*, 2(1).
- Marinoiu, C. (2018). Average Monthly Rainfall Forecast in Romania by Using K-Nearest Neighbors Regression. *Annals of the „Constantin Brâncuși” University of Târgu Jiu*, 4.
- Seruni, Sekar, D., Furqon, Tanzil, M., Wihandika, Cahya, R. (2020). Sistem Prediksi Pertumbuhan Jumlah Penduduk Kota Malang Menggunakan Metode K-Nearest Neighbor Regression. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 4(4).
- Diajeng, S. S., Muhammad, T.F. & Randy, C.W. (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.
- Sindhoora, K. M., Spandana, K. U., Ivanov, D., Borisova, E., Raghavendra, U., Rai, S., Kabekkodu, S. P., Mahato, K. K., Mazumder, N. (2021). Machine-learning based classification of Stokes-Mueller polarization images for tissue characterization. *Journal of Physics: Conference Stories*. 1859.