

## Penerapan Data Mining untuk Menganalisis Faktor-Faktor yang Mempengaruhi Terhadap Kenaikan Harga Minyak Goreng

Putri Handayani<sup>1</sup>, Tri Susanto<sup>2</sup>, Aurino Djamaris<sup>3</sup>, Mirsa Diah Novianti<sup>4</sup>

<sup>1,2,4</sup> Teknik Industri, FTIK, Universitas Bakrie, Jakarta, Indonesia

<sup>1</sup>[putrihandayani825@gmail.com](mailto:putrihandayani825@gmail.com) <sup>2</sup>[tri.susanto@bakrie.ac.id](mailto:tri.susanto@bakrie.ac.id) <sup>4</sup>[mirsanovaianti@bakrie.ac.id](mailto:mirsanovaianti@bakrie.ac.id)

<sup>3</sup> Manajemen, FEIS, Jakarta, Indonesia

<sup>3</sup>[aurino@bakrie.ac.id](mailto:aurino@bakrie.ac.id)

Submitted : 2023-01-09 | Reviewed : 2023-01-19 | Accepted : 2023-01-25

**Abstract-** *Cooking oil is currently one type of oil that is commonly used by the people of Indonesia. The price of cooking oil often changes, so that fluctuations in the price of cooking oil become very difficult to predict due to unstable movements. The limited domestic cooking oil supply has resulted in high domestic cooking oil prices. This research was carried out based on these problems by applying data mining to obtain relationships or patterns that might provide useful indications. The data mining process uses the Cross-Industry Standard Process for Data Mining (CRISP-DM) method with multiple linear regression algorithms to collect data for 5 years from January 2017 to April 2022, obtained through websites or official Instagram accounts, from relevant data and accessed via the Internet. A multiple linear regression technique was used to determine the effect of five independent variables, namely CPO production, cooking oil production, total domestic cooking oil consumption, total cooking oil exports, and world CPO prices on variable Y, namely the average price of domestic cooking oil. The results showed that five variables affected 63.3% of variable Y (the average price of domestic cooking oil), while the remaining 36.7% was influenced by other factors that were not calculated in the study.*

**Keywords—** *Cooking Oil; Data Mining; Multiple Linear Regression.*

**Abstrak—** Minyak goreng saat ini merupakan salah satu jenis minyak yang umum digunakan oleh masyarakat Indonesia. Harga dari minyak goreng sering mengalami perubahan sehingga fluktuasi dari harga minyak goreng menjadi sangat sulit diprediksi karena pergerakan yang tidak stabil. Keterbatasan stok minyak goreng dalam negeri mengakibatkan tingginya harga minyak goreng domestik. Berdasarkan permasalahan tersebut, maka penelitian ini dilakukan dengan penerapan *data mining* untuk mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat. Proses data mining menggunakan metode *Cross-Industry Standard Process for Data Mining (CRISP-DM)* dengan algoritma regresi linear berganda dengan menghimpun data selama 5 tahun dari Januari 2017 sampai dengan April 2022 yang diperoleh melalui *website* resmi atau *official* Instagram dari data relevan dan diakses melalui internet. Teknik regresi linear berganda yang digunakan untuk mengetahui pengaruh antara 5 variabel bebas, yaitu produksi CPO, produksi minyak goreng, total konsumsi minyak goreng domestik, total ekspor minyak goreng, harga CPO dunia terhadap variabel Y yaitu harga rata-rata minyak goreng domestik. Hasil penelitian menunjukkan bahwa 5 variabel mempengaruhi 63.3% terhadap variabel Y (harga rata-rata minyak goreng domestik) sedangkan sisanya sebesar 36.7% dipengaruhi oleh faktor lain yang tidak dihitung dalam penelitian.

**Kata kunci:** Minyak Goreng; Data Mining; Regresi Linear Berganda.

## PENDAHULUAN

Meningkatnya kebutuhan masyarakat Indonesia terhadap produk minyak goreng kemasan tentunya harus diikuti dengan *supply* atau ketersediaan minyak goreng kemasan yang memadai sehingga *demand* dari minyak goreng kemasan di dalam negeri dapat tercukupi. Kenaikan harga minyak goreng kemasan tersebut dapat menurunkan tingkat pembelian produk minyak goreng di kalangan masyarakat. Sementara itu, data menunjukkan bahwa tingkat konsumsi minyak goreng terus mengalami kenaikan. Berdasarkan data Pusat Informasi Harga Pangan Strategis Nasional komoditas minyak goreng kemasan per kilogram (Kg) pada April 2022 mencapai Rp. 26.550. Melonjaknya harga minyak goreng kemasan di Indonesia di tahun 2022 ini memunculkan beberapa spekulasi atau dugaan yang menjadi faktor pengaruh naiknya harga minyak goreng domestik di tahun 2022 tersebut.

Khoiru Rizqy Rambe dan Nunung Kusnadi (2018) menyebutkan bahwa faktor yang mempengaruhi kenaikan harga minyak goreng dikarenakan adanya permintaan yang berlebih dibanding penawarannya. Selain itu, harga minyak goreng domestik juga dapat lebih stabil jika ketersediaan minyak goreng dapat terjamin salah satunya dilihat dari sisi produksi *Crude Palm Oil (CPO)*. Tingginya produksi *CPO* akan mempengaruhi tingkat produksi minyak goreng domestik dikarenakan pemanfaatan *CPO* sebagai bahan baku utama pembuatan minyak goreng. Hastuti, Zulaicha Dwi, dkk. (2015) menyebutkan bahwa persentase dari produksi *CPO* yang dijadikan sebagai minyak goreng adalah sebesar 79%, industri kimia 14%, industri sabun 4% dan industri margarin 3%. Sedangkan menurut Siregar (2014), harga minyak goreng di Indonesia dipengaruhi oleh harga *Crude Palm Oil (CPO)* Dunia. Peningkatan harga *CPO* akan memicu ekspor yang berlebihan yang menyebabkan keterbatasan produk minyak goreng di dalam negeri sehingga harga minyak goreng domestik sebagai salah satu produk turunan *CPO* ikut meningkat.

Untuk mengetahui lebih lanjut terkait pengaruh faktor-faktor yang mempengaruhi kenaikan harga minyak goreng yang terjadi di tahun 2022 maka perlu dilakukan analisis lebih lanjut dari berbagai spekulasi atau faktor yang ada apakah produksi *Crude Palm Oil/CPO*, produksi minyak goreng domestik, total konsumsi minyak goreng domestik, jumlah ekspor minyak goreng, dan harga *Crude Palm Oil* dunia tersebut berpengaruh secara positif perubahan harga minyak goreng dan analisis menggunakan proses data mining dapat memperoleh informasi atau output knowledge yang akan berguna untuk memahami permasalahan terhadap keterbatasan minyak goreng di pasaran domestik.

Dalam analisis ini memanfaatkan database menggunakan teknik data mining dengan algoritma regresi. Dimana, database dalam penelitian ini merupakan data-data relevan yang dihimpun oleh peneliti mulai dari jumlah produksi *CPO*, produksi minyak goreng domestik, jumlah konsumsi minyak goreng domestik, jumlah ekspor minyak goreng, harga rata-rata minyak goreng, harga *CPO* dunia dan lain-lain. Database yang digunakan kajian ini merupakan data *time series* selama 5 tahun mulai dari tahun 2017 hingga 2022. Hal ini dikarenakan agar data tersebut lebih mudah untuk dipahami sehingga fluktuasi dari data dapat tergambarkan.

## TINJAUAN PUSTAKA

### A. Minyak goreng

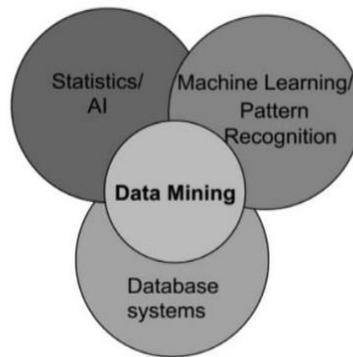
Minyak goreng adalah bahan pangan dengan komposisi utama trigliserida yang berasal dari bahan nabati dengan tanpa perubahan kimiawi termasuk hidrogenasi, pendinginan dan telah melalui proses rafinasi atau pemurnian yang digunakan untuk menggoreng (Risti, 2016). Minyak goreng yang umum digunakan oleh masyarakat Indonesia salah satunya adalah jenis minyak goreng kelapa sawit. Hingga saat ini kegiatan menggoreng menggunakan minyak goreng jenis kelapa sawit masih banyak digunakan dibandingkan dengan jenis minyak goreng lain karena harganya yang terjangkau. Minyak dan lemak memiliki titik didih yang tinggi (sekitar 200° C) sehingga minyak dan lemak dapat digunakan untuk menggoreng sehingga makanan yang digoreng akan kehilangan sebagian kadar airnya dan menjadi lebih kering serta memberikan rasa gurih.

Menurut Ketaren (2012), jenis-jenis minyak goreng diantaranya minyak kelapa sawit, minyak jagung, minyak wijen, minyak kedelai, minyak kemiri, minyak kelapa, minyak biji bunga matahari, minyak alpukat, minyak kacang tanah, minyak jarak.

## B. Data Mining

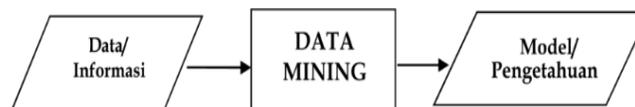
Data merupakan sebuah rekaman dari fakta-fakta, konsep, atau instruksi pada media penyimpanan untuk pemrosesan dengan cara otomatis dan presentasi sebagai informasi yang dapat dimengerti oleh manusia (Inmon, 2005). Sedangkan big data adalah istilah yang menggambarkan volume data dalam jumlah yang besar baik terstruktur maupun tidak terstruktur yang disimpan dan dapat dianalisis yang menghasilkan suatu informasi untuk mendukung pengambilan keputusan pada suatu organisasi atau bisnis

Data mining merupakan suatu proses otomatis terhadap data yang sudah ada untuk mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat (Kusrini dan Emha Taufiq Luthfi, 2009). Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakit dari berbagai database besar (Turban, dkk. 2005). Tan dalam (Arhami, Muhammad dan Muhammad Nasir, 2020) menyatakan bahwa data mining merupakan irisan dari *machine learning*, *artificial intelligence*, pengenalan pola, *statistic*, dan sistem basis data.



GAMBAR 1. DATA MINING IRISAN DARI BEBERAPA BIDANG

Maka dari beberapa definisi data mining menurut para ahli di atas dapat disimpulkan bahwa data mining merupakan suatu proses ekstraksi data besar yang digunakan untuk mendapatkan model atau informasi dari suatu permasalahan yang akan dianalisis. Data mining memiliki tujuan untuk menemukan pola yang sebelumnya tidak diketahui untuk menyelesaikan permasalahan-permasalahan yang menjadi tujuan dilakukannya data mining. Melalui data mining perusahaan atau seorang analis/peneliti dapat menemukan informasi dalam data yang begitu besar melalui pengolahan data yang tepat dan efektif dengan berbagai macam metode yang ada di dalam data mining.

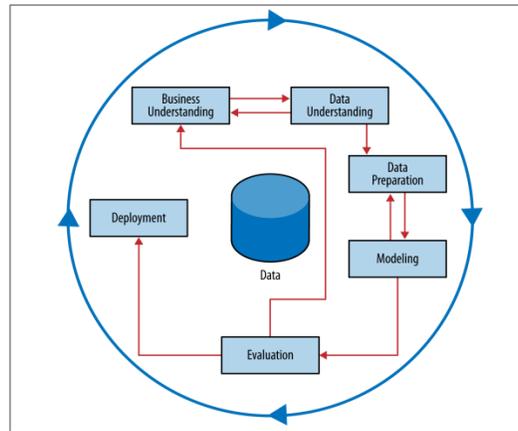


GAMBAR 2 INPUT PROSES DAN OUTPUT DATA MINING

Sumber: Jose Hernandez-Orallo, *Introduction to Data Mining*

## C. CROSS-INDUSTRY STANDARD PROCESS FOR DATA MINING (CRISP-DM)

*CRISP-DM* merupakan salah satu model standar proses data mining komprehensif yang dikembangkan pada tahun 1996 yang digunakan sebagai strategi pemecahan masalah secara umum dari bisnis atau unit penelitian dan dapat digunakan mulai dari pemula hingga kepada ahli data mining. Model *CRISP-DM* ini memiliki 6 tahapan yang memberikan gambaran tentang siklus hidup proyek data mining menurut (Foster Provost & Tom Fawcett, 2013)



GAMBAR 3 TAHAPAN *CRISP-DM*  
Sumber: Provost, Foster and Tom Fawcett (2013)

### 1. *Business Understanding*

Tahapan pada pemahaman bisnis merupakan bagian yang berperan besar tergantung kepada kreativitas dari analis atau peneliti karena sangat penting untuk memahami masalah yang akan dipecahkan dari suatu data agar menjadi suatu pemecahan masalah yang jelas dan tidak ambigu.

### 2. *Data Understanding*

Dalam pemahaman data seorang analis perlu melihat dan menggali lebih dalam untuk memecahkan masalah bisnis atau unit penelitian dari data yang ada untuk kemudian mencocokkan data tersebut karena tidak semua data yang ada relevan dengan masalah atau tujuan dari unit penelitian itu sendiri..

### 3. *Data Preparation*

Tahap persiapan data sering dilakukan bersamaan dengan proses pemahaman data. Dimana seorang analis atau peneliti bisa diubah atau ditambahkan sehingga menghasilkan data yang lebih baik sebelum dilanjutkan ke tahap selanjutnya.

### 4. *Modeling*

Tahap pemodelan ini dilakukan untuk mendapatkan output semacam model atau pola yang ada pada data. Proses modelling dapat kembali ke fase persiapan data jika peneliti merasa memerlukan hal tersebut untuk menjadikan data ke dalam bentuk yang sesuai dengan spesifikasi kebutuhan teknik data mining tertentu.

### 5. *Evaluation*

Tahap evaluasi bertujuan untuk menilai hasil data mining secara keseluruhan untuk memastikan bahwa hasil model yang telah dilakukan valid dan dapat diandalkan sebelum dilanjutkan dan disebarakan untuk digunakan.

### 6. *Deployment*

Tahap penyebaran adalah proses menggunakan model yang telah dihasilkan. Contoh sederhana dari tahap penyebaran adalah pembuatan laporan.

## METODE PENELITIAN

### Pendekatan Penelitian

Metode yang digunakan dalam penelitian ini adalah metode kuantitatif dengan pendekatan deskriptif. Dimana metode kuantitatif dalam penelitian digunakan untuk menganalisis pengaruh jumlah ekspor minyak goreng terhadap kenaikan harga minyak goreng di Indonesia yang terjadi pada awal tahun 2022 dengan menggunakan teknik data mining dengan metode *Cross-Industry Standard Process for Data Mining (CRISP-DM)*. Sedangkan metode deskriptif digunakan untuk menampilkan visualisasi dan menggambarkan pola dan kecenderungan yang ada di dalam data secara sistematis sehingga lebih mudah dipahami.

## Jenis dan Sumber Data

Penelitian ini menggunakan data sekunder dengan jenis data runtun waktu (*time series*) selama kurun waktu 2017-2022. Data-data relevan yang dihimpun oleh peneliti mulai dari jumlah produksi, jumlah konsumsi minyak goreng domestik, *demand* minyak sawit dunia, persentase alokasi turunan CPO, jumlah ekspor CPO dan salah satu turunannya yaitu minyak goreng, harga eceran rata-rata minyak goreng, harga CPO dan lain-lain. Data yang diperlukan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik, USDA (*United States Department of Agriculture*), Kemenperin, Infopangan.go.id, Gabungan Pengusaha Kelapa Sawit Indonesia (GAPKI) atau *Indonesian Palm Oil Association* (IPOA), Bea cukai (beacukai.go.id), Kementerian Perdagangan Republik Indonesia (kemendag.go.id), Kementerian Pertanian Republik Indonesia, Pusat Informasi Harga Pangan Strategis Nasional (PIHPS Nasional), *World Bank Commodity Price Data (The Pink Sheet)*.

## Metode Pengumpulan Data

Metode pengumpulan data dilakukan melalui website-website resmi atau Instagram seperti Badan Pusat Statistik, USDA, Kemenperin, Infopangan.go.id, Gabungan Pengusaha Kelapa Sawit Indonesia (GAPKI) atau *Indonesian Palm Oil Association* (IPOA) (Instagram @gapki\_ipoa), Bea cukai (beacukai.go.id), Kementerian Perdagangan Republik Indonesia (kemendag.go.id), dan Kementerian Pertanian Republik Indonesia

## Metode Analisis Data

Metode analisis data yang dilakukan dalam penelitian ini menggunakan teknik data mining dengan metode *Cross-Industry Standard Process for Data Mining (CRISP-DM)* untuk digunakan sebagai strategi pemecahan masalah secara umum dari bisnis atau unit penelitian dan dapat digunakan mulai dari pemula hingga kepada ahli data mining.

Dalam penelitian ini, peneliti menggunakan Aplikasi Orange untuk melakukan proses pengolahan data mulai dari proses *cleaning data*, hingga kepada pembuatan model yang akan menghasilkan suatu informasi atau pengetahuan berdasarkan proses data mining yang telah dilakukan.

## HASIL DAN PEMBAHASAN

### A. Himpunan Data

Beberapa variabel yang digunakan dalam penelitian diantaranya yaitu:

1. Variabel X1 dalam penelitian ini adalah jumlah produksi *Crude Palm Oil (CPO)* yang diukur dengan satuan unit ton.
2. Variabel X2 dalam penelitian ini adalah jumlah produksi minyak goreng yang diukur dengan satuan unit ton.
3. Variabel X3 dalam penelitian ini adalah total konsumsi minyak goreng domestik yang diukur dengan satuan unit.
4. Variabel X4 dalam penelitian ini adalah jumlah ekspor minyak goreng yang diukur dengan satuan unit ton.
5. Variabel X5 dalam penelitian ini adalah harga *Crude Palm Oil (CPO)* Dunia yang diukur dengan satuan unit dollar per metrik ton.
6. Variabel Y dalam penelitian ini adalah harga minyak goreng domestik yang diukur dengan satuan unit rupiah per Kg.

Dimana, variabel produksi CPO, produksi minyak goreng, total konsumsi domestik, ekspor minyak goreng, harga CPO dunia, harga minyak goreng domestik bernilai skala interval. Dapat dilihat pada Tabel 1 merupakan contoh himpunan data yang akan digunakan dalam penelitian.

Tabel 1 Contoh Himpunan Data Minyak Goreng

Periode	Produksi CPO (Ribu Ton)	Produksi Minyak Goreng (79%) (Ribu Ton)	Total Konsumsi Domestik	Ekspor Olahan CPO/Minyak Goreng (Ribu Ton)	Harga CPO Dunia (\$/MT)	Harga Rata-Rata Minyak Goreng di Indonesia (Per Kg)
Jul-17	3,013	2,380	821	1,577	\$ 720.48	Rp 15,050
Aug-17	3,037	2,399	937	2,051	\$ 717.95	Rp 15,050
Sep-17	3,235	2,555	978	1,780	\$ 754.25	Rp 15,050
Oct-17	2,942	2,324	985	1,682	\$ 746.79	Rp 14,950
Nov-17	2,851	2,253	931	1,725	\$ 728.86	Rp 14,950
Dec-17	2,848	2,250	946	1,728	\$ 679.17	Rp 15,000
Jan-18	2,739	2,164	974	2,303	\$ 703.45	Rp 14,950
Feb-18	2,538	2,005	902	2,119	\$ 709.44	Rp 14,900

Tingginya produksi *CPO* akan mempengaruhi tingkat produksi minyak goreng domestik dikarenakan pemanfaatan *CPO* sebagai bahan baku utama pembuatan minyak goreng. Hastuti, Zulaicha Dwi, dkk. (2015) menyebutkan bahwa persentase dari produksi *CPO* yang dijadikan sebagai minyak goreng adalah sebesar 79%, industri kimia 14%, industri sabun 4% dan industri margarin 3%. Sedangkan menurut Siregar (2014), harga minyak goreng di Indonesia dipengaruhi oleh harga *Crude Palm Oil (CPO)* Dunia. Peningkatan harga *CPO* akan memicu ekspor yang berlebihan yang menyebabkan keterbatasan produk minyak goreng di dalam negeri sehingga harga minyak goreng domestik sebagai salah satu produk turunan *CPO* ikut meningkat.

## B. Pre-Processing atau Cleaning Data

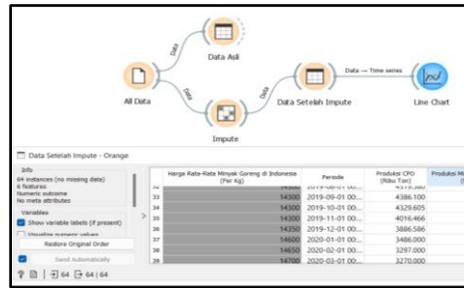
### a. Mengisi Missing Value Pada Variabel Data

Dari beberapa variabel yang ada pada himpunan data penelitian, terdapat nilai variabel yang kosong seperti yang terlihat pada Tabel 2 yang merupakan contoh *missing value* dari variabel pada himpunan data penelitian ini.

Tabel 2 Contoh Missing Value

Periode	Produksi CPO (Ribu Ton)	Produksi Minyak Goreng (79%) (Ribu Ton)	Total Konsumsi Domestik	Ekspor Olahan CPO/Minyak Goreng (Ribu Ton)	Harga CPO Dunia (\$/MT)	Harga Rata-Rata Minyak Goreng di Indonesia (Per Kg)
Jul-19	4,288	3,387	1437	2,239	\$ 543.88	Rp 14,300
Aug-19	4,320	3,412	1512	2,327	\$ 586.12	Rp 14,300
Sep-19	4,386	3,465	1410	2,569	\$ 580.30	Rp 14,300
Oct-19	4,330	3,420		2,428	\$ 591.35	Rp 14,300
Nov-19	4,016	3,173		2,566	\$ 685.41	Rp 14,300
Dec-19	3,887	3,070		2,603	\$ 763.73	Rp 14,350
Jan-20	3,486	2,754	1494	1,246	\$ 810.07	Rp 14,600
Feb-20	3,297	2,605	1547	1,661	\$ 728.81	Rp 14,650

Maka dari itu, perlu melakukan manipulasi terhadap data yang kosong dengan menggunakan *model-based imputer (simple tree)* pada aplikasi Orange. Model ini digunakan untuk membuat model prediksi dari *missing value* tersebut berdasarkan *values* dari variabel-variabel lainnya. Dapat dilihat *model-based imputer (simple tree)* seperti pada Gambar 3 yang merupakan hasil model



GAMBAR 4 HASIL MODEL-BASED IMPUTER (SIMPLE TREE) UNTUK VARIABEL MISSING VALUES

**b. Menyesuaikan Spesifikasi Periode Data Penelitian**

Penelitian ini menggunakan periode data dari tahun 2017 sampai dengan tahun 2022. Dimana himpunan data penelitian ini disesuaikan berdasarkan laporan bulanan. Oleh karena itu, perlu melakukan beberapa konversi data karena beberapa bentuk laporan variabel ada yang berbentuk laporan harian, laporan mingguan sehingga peneliti melakukan konversi data untuk menyesuaikan menjadi bentuk laporan per bulan. Selain itu, terdapat data harga minyak goreng domestik per provinsi oleh karena itu perlu melakukan konversi data atau mengambil data yang diperlukan yaitu data harga rata-rata minyak goreng domestik di Indonesia Dapat dilihat pada Tabel 4 yang merupakan contoh konversi data dari laporan per provinsi menjadi laporan per bulan dengan data harga rata-rata minyak goreng domestik.

Tabel 4: Contoh Tipe data Pada variabel

Perkembangan Harga Pangan					
Periode		Jan 2017 - Sep 2022			
Komoditas		Minyak Goreng Kemasan Bermerk 1 (kg)			
Tipe Laporan		Laporan Bulanan			
No.	Provinsi	07/2017	08/2017	09/2017	
I	SEMUA PROVINSI	Rp 15,050	Rp 15,050	Rp 15,050	Rp 15,050
II	ACEH	Rp 15,500	Rp 15,500	Rp 15,500	Rp 15,500
III	SUMATERA UTARA	Rp 14,300	Rp 14,300	Rp 14,300	Rp 14,300
IV	SUMATERA BARAT	Rp 15,000	Rp 15,000	Rp 15,000	Rp 15,000
V	RIAU	Rp 15,300	Rp 15,300	Rp 15,300	Rp 15,250
VI	KEPULAUAN RIAU	Rp 13,650	Rp 13,650	Rp 13,650	Rp 13,400
VII	JAMBI	Rp 14,350	Rp 14,350	Rp 13,800	Rp 13,800
VIII	BENGKULU	Rp 14,250	Rp 14,000	Rp 13,750	Rp 13,750

**c. Merubah Tipe Data Pada Variabel**

Perubahan tipe data dari himpunan data penelitian dilakukan agar dapat dilanjutkan ke proses selanjutnya atau proses perhitungan/modelling. Sehingga perubahan tipe data pada variabel periode yang bertipe *datetime* akan diubah menjadi *timestamp* agar data bernilai numerik.

Table 3. Variabel Periode Bernilai Timestamp

Periode (Timestamp)	Produksi CPO (Ribu Ton)	Produksi Minyak Goreng (79%) (Ribu Ton)	Total Konsumsi Domestik	Ekspor Olahahan CPO/Minyak Goreng (Ribu Ton)	Harga CPO Dunia (\$/MT)	Harga Rata-Rata Minyak Goreng di Indonesia (Per Kg)
1498867200	3,013	2,380	821	1,577	\$ 720.48	Rp 15,050
1501545600	3,037	2,399	937	2,051	\$ 717.95	Rp 15,050
1504224000	3,235	2,555	978	1,780	\$ 754.25	Rp 15,050
1506816000	2,942	2,324	985	1,682	\$ 746.79	Rp 14,950
1509494400	2,851	2,253	931	1,725	\$ 728.86	Rp 14,950
1512086400	2,848	2,250	946	1,728	\$ 679.17	Rp 15,000
1514764800	2,739	2,164	974	2,303	\$ 703.45	Rp 14,950

#### d. Transformasi Data

Dari himpunan data yang telah dihimpun memiliki satuan ukur yang berbeda sehingga dilakukan proses transformasi data pada himpunan data asli sebelum dilanjutkan ke proses selanjutnya. Proses transformasi data yang digunakan dalam penelitian ini yaitu Logaritma Natural (Ln) yang bertujuan agar hasil dari model regresi linear yang dilakukan tidak bias. Oleh karena itu, perlu melakukan proses transformasi data menggunakan Microsoft Excel seperti pada Tabel 5 yang merupakan proses transformasi data asli menjadi bentuk Logaritma Natural (Ln).

Tabel 5. Contoh Transformasi Data Menjadi Data Logaritma Natural

Ln Periode	Ln Produksi CPO (Ribu Ton)	Ln Produksi Minyak Goreng	Ln Total Konsumsi Domestik	Ln Ekspor Olahahan CPO/Minyak Goreng (Ribu Ton)	Ln Harga CPO Dunia (\$/MT)	Ln Harga Rata-Rata Minyak Goreng di Indonesia (Per Kg)
21.117	8.005	7.769	6.817	7.432	6.715	9.608
21.119	7.878	7.642	6.807	7.433	6.695	9.608
21.121	7.938	7.702	6.825	7.502	6.658	9.608
21.123	7.912	7.677	6.818	7.338	6.623	9.608
21.124	7.928	7.692	6.824	7.407	6.637	9.608
21.126	7.865	7.629	6.789	7.207	6.600	9.608
21.128	8.011	7.775	6.711	7.363	6.580	9.619
21.130	8.019	7.783	6.843	7.626	6.576	9.619
21.132	8.082	7.846	6.886	7.484	6.626	9.619
21.133	7.987	7.751	6.893	7.428	6.616	9.612

### C. Penerapan Metode *Cross-Industry Standard Process for Data Mining (CRISP-DM)*

#### *Business Understanding*

Proses *business understanding* dilakukan oleh penulis dengan memahami masalah yang akan dipecahkan dari suatu data agar menjadi suatu pemecahan masalah yang jelas dan tidak ambigu. Dimana dalam penelitian ini masalah yang akan dipecahkan terkait faktor-faktor penyebab dari terjadinya kenaikan harga minyak goreng kemasan domestik pada tahun 2022 di Indonesia. Peneliti melakukan analisa pada beberapa variabel seperti produksi *Crude Palm Oil/CPO*, produksi minyak goreng, total konsumsi domestik, jumlah minyak goreng yang di ekspor ke luar negeri dan harga *Crude Palm Oil/CPO* dunia, sehingga pada penelitian ini perlu menganalisis berdasarkan fakta yang ada pada data apakah hal tersebut benar bahwa faktor-faktor yang menjadi variabel dalam penelitian ini akan mempengaruhi harga minyak goreng kemasan domestik dilihat melalui perspektif *supply and demand* dengan penerapan proses data mining teknik regresi.

#### *Data Understanding*

Proses *data understanding* dalam penelitian ini penulis memahami lebih dalam masalah yang akan dipecahkan dan mencocokkan data-data yang relevan. Dimana dalam penelitian ini masalah yang akan dipecahkan adalah faktor-faktor penyebab dari terjadinya kenaikan harga minyak goreng kemasan domestik pada tahun 2022 di Indonesia. Dengan menganalisis berdasarkan fakta yang ada pada data apakah hal tersebut benar bahwa variabel-variabel bebas yang digunakan dalam penelitian akan mempengaruhi tinggi rendahnya harga minyak goreng kemasan domestik. Sehingga pada proses data understanding ini peneliti memahami data-data relevan dari masalah seperti data produksi *Crude Palm Oil/CPO*, persentase alokasi *CPO* menjadi minyak goreng, produksi minyak goreng domestik, total konsumsi minyak goreng domestik, total ekspor minyak goreng, harga *Crude Palm Oil/CPO dunia*, dan harga rata-rata minyak goreng domestic.

#### *Data Preparation*

Proses *data preparation* dalam penelitian ini perlu melakukan pengumpulan data sehingga menjadi suatu kumpulan data-data yang dijadikan database dalam penelitian. Dalam proses ini peneliti menghimpun data persentase alokasi *Crude Palm Oil (CPO)* yang diolah menjadi minyak goreng. Dapat dilihat pada tabel 6 yang merupakan data persentase alokasi *Crude Palm Oil (CPO)* menjadi minyak goreng.

Tabel 6 Persentase Alokasi CPO Menjadi Minyak Goreng

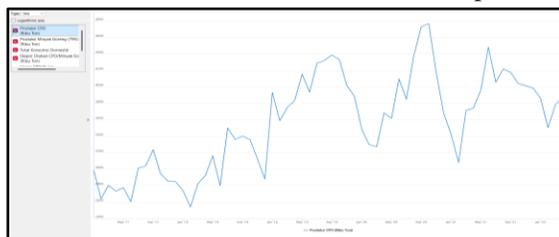
Persentase Alokasi CPO di Indonesia (%)	
Minyak goreng	79
Margarin	3
Sabun	4
Bidang Kimia	14
<b>Total</b>	<b>100</b>

Sumber: Hastuti, Zulaicha Dwi, dkk. (2015)

Selain itu, perlu menghimpun data-data variabel yang akan diolah ke proses selanjutnya. Adapun data-data variabel yang digunakan seperti:

1. Data Jumlah Produksi *Crude Palm Oil (CPO)*

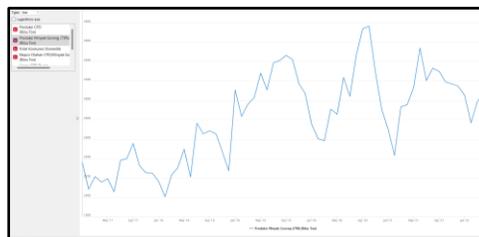
Dapat dilihat pada Gambar 5, grafik data jumlah produksi *Crude Palm Oil (CPO)* di Indonesia yang dihimpun oleh peneliti dari Badan Pusat Statistik dan GAPKI mulai dari Januari 2017 sampai dengan April 2022.



GAMBAR 5 GRAFIK DATA JUMLAH PRODUKSI CPO JANUARI 2017 S.D. APRIL 2022

2. Data Jumlah Produksi Minyak Goreng

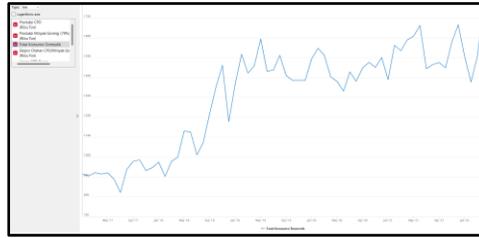
Dapat dilihat pada Gambar 4.7, Grafik data jumlah produksi minyak goreng yang diperoleh dari 79% alokasi produksi *Crude Palm Oil (CPO)* dan dihimpun oleh peneliti dari Badan Pusat Statistik dan GAPKI mulai dari Januari 2017 sampai dengan April 2022.



GAMBAR 6 GRAFIK DATA JUMLAH PRODUKSI MINYAK GORENG JANUARI 2017 S.D. APRIL 2022

3. Data Total Konsumsi Minyak Goreng Domestik

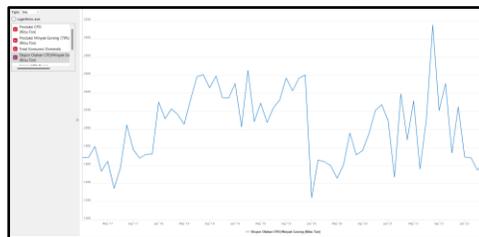
Gambar 6 menunjukkan data konsumsi minyak goreng domestik Januari 2017 sampai dengan Maret 2022 yang peneliti himpun dari GAPKI (Gabungan Pengusaha Kelapa Sawit Indonesia) melalui *website* resmi dan *official* Instagram GAPKI.



GAMBAR 7 GRAFIK DATA TOTAL KONSUMSI MINYAK GORENG DOMESTIK JANUARI 2017 S.D. APRIL 2022

#### 4. Data Jumlah Ekspor Minyak Goreng

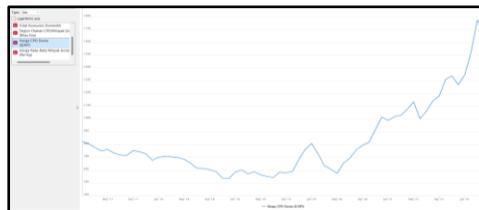
Data terkait jumlah ekspor olahan *crude palm oil (CPO)*/minyak goreng ini penulis melakukan rekapitulasi data-data dari berbagai sumber seperti Badan Pusat Statistik dan *website* ataupun *official* Instagram Gabungan Pengusaha Kelapa Sawit Indonesia (GAPKI). Dapat dilihat pada Gambar 8 yang menunjukkan grafik data jumlah ekspor olahan *Crude Palm Oil (CPO)*/minyak goreng mulai dari Januari 2020 sampai dengan Maret 2022.



GAMBAR 8 GRAFIK DATA JUMLAH EKSPOR MINYAK GORENG JANUARI 2017 S.D. APRIL 2022

#### 5. Data Harga *Crude Palm Oil (CPO)* Dunia

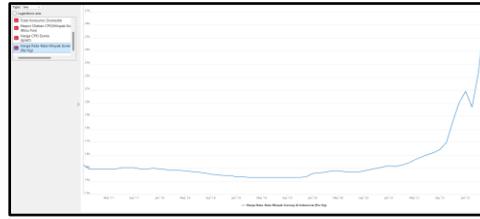
Data terkait harga *Crude Palm Oil (CPO)* Dunia mulai dari Januari 2017 sampai dengan Februari 2022. Data harga *Crude Palm Oil (CPO)* Dunia dihimpun oleh peneliti dari *website World Bank Commodity Price Data (The Pink Sheet)* yang diakses melalui internet untuk digunakan sebagai bahan analisis selanjutnya. Dapat dilihat pada Gambar 9 merupakan grafik data harga *Crude Palm Oil (CPO)* Dunia.



GAMBAR 9 GRAFIK DATA HARGA CRUDE PALM OIL (CPO) JANUARI 2017 S.D. APRIL 2022

#### 6. Data Harga Minyak Goreng Domestik

Data terkait harga minyak goreng kemasan domestik di Indonesia mulai dari Januari 2017 sampai dengan Februari 2022. Data harga minyak goreng kemasan dihimpun dari *website resmi Pusat Informasi Harga Pangan Strategis Nasional (PIHPS Nasional)* yang diakses melalui internet. Dapat dilihat pada Gambar 10 merupakan grafik data harga minyak goreng domestik.



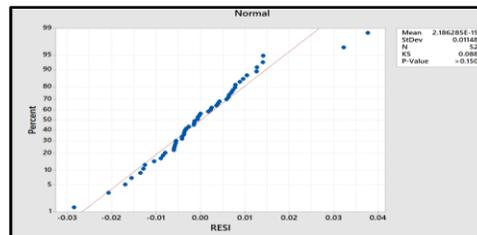
GAMBAR 10 GRAFIK DATA HARGA MINYAK GORENG DOMESTIK JANUARI 2017 S.D. APRIL 2022

## Modelling

### a. Uji Asumsi Klasik

#### 1. Uji Normalitas

Uji asumsi klasik yang pertama yaitu uji normalitas untuk melihat apakah data-data yang sudah melalui tahapan *cleaning data* atau *pre-processing* sudah berdistribusi normal. Pengujian normalitas ini menjadi salah satu syarat sebelum dilakukan uji model analisis regresi. Dapat dilihat pada Gambar 11 yang menunjukkan grafik normalitas residual bahwa data tersebar di sekitar garis lurus dan memiliki nilai p-value  $>0.150$ .



GAMBAR 11 GRAFIK NORMALITAS RESIDUAL

#### 2. Uji Multikolinearitas

Selanjutnya, uji asumsi klasik yang dilakukan dalam penelitian ini adalah uji multikolinearitas. Uji multikolinearitas disini bertujuan untuk menguji apakah di antara variabel bebas terjadi korelasi. Uji multikolinearitas yang digunakan yaitu *Variance Inflation Factor (VIF)*. Jika nilai VIF lebih besar dari 10 hal tersebut mengindikasikan bahwa terjadi multikolinearitas begitupun sebaliknya jika nilai VIF lebih kecil dari 10 hal tersebut mengindikasikan bahwa tidak terjadi multikolinearitas dan hasilnya sebagai berikut:

$$\text{Tolerance} = 1 - r^2 \text{ dan } \text{VIF} = 1/\text{Tolerance}$$

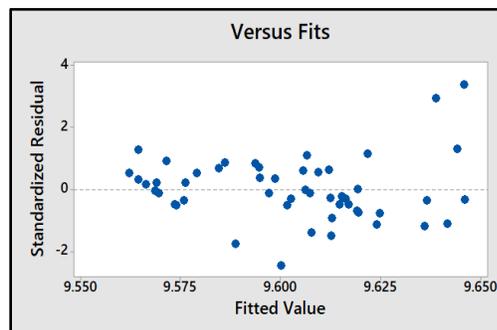
Pada Tabel 7 yang merupakan hasil nilai VIF antara variabel bebas. Dimana hasil tersebut menunjukkan bahwa pada variabel X1 (produksi *Crude Palm Oil*) dan Variabel X2 (produksi minyak goreng) terjadi korelasi sempurna dengan nilai  $r = 1$ . Maka dari itu, penulis hanya menggunakan salah satu variabel bebas yang mengalami korelasi sempurna diantara variabel X1 dan X2. Selanjutnya hanya akan meneliti menggunakan salah satu variabel yang terjadi korelasi sempurna yaitu penggunaan variabel X1 (produksi *CPO*) untuk proses selanjutnya. Sedangkan untuk variabel bebas lainnya tidak terjadi multikolinearitas karena nilai VIF lebih kecil dari 10.

Table 7 Hasil Uji Multikolinearitas

	<i>r</i>	<i>r</i> <sup>2</sup>	Tolerance	VIF
<i>rx1x2</i>	1	1	0	
<i>rx1x3</i>	0.81116	0.6580	0.3420	2.9239
<i>rx1x4</i>	0.20580	0.0424	0.9576	1.0442
<i>rx1x5</i>	0.04806	0.0023	0.9977	1.0023
<i>rx2x3</i>	0.81116	0.6580	0.3420	2.9238
<i>rx2x4</i>	0.20579	0.0424	0.9576	1.0442
<i>rx2x5</i>	0.04806	0.0023	0.9977	1.0023
<i>rx3x4</i>	0.24686	0.0609	0.9391	1.0649
<i>rx3x5</i>	0.05353	0.0029	0.9971	1.0029
<i>rx4x5</i>	-0.23337	0.0545	0.9455	1.0576

### 3. Uji Heteroskedastisitas

Uji asumsi klasik selanjutnya yaitu uji heteroskedastisitas, dimana uji heteroskedastisitas dalam penelitian ini menggunakan uji grafik plot sehingga jika penyebaran titik-titik yang terbentuk secara acak atau tidak membentuk suatu pola tertentu. Maka dapat dilihat pada Gambar 12 menunjukkan bahwa dilihat dari grafik versus fits menunjukkan bahwa data tersebar di sumbu x dan y maka hal tersebut mengindikasikan tidak terjadi heteroskedastisitas.



GAMBAR 12 HASIL Uji HETEROSKEDASITAS

#### b. Penerapan Metode Analisis Regresi

##### *Pearson Correlation*

Karena data dalam penelitian ini merupakan data interval atau rasio maka digunakan *Pearson correlation* untuk melihat hubungan linier antara dua variabel antara variabel X dan variabel Y. Hubungan positif dan negatif pada koefisien variabel menunjukkan arah hubungan variabel. Dimana jika koefisien hubungannya positif artinya hubungan antar variabelnya berbanding lurus atau searah, jika nilai variabel x semakin tinggi maka nilai variabel y juga semakin tinggi. Jika koefisien hubungan negatif artinya hubungan antar variabelnya berbanding terbalik atau tidak searah, jika nilai variabel x semakin tinggi maka nilai variabel y semakin rendah begitupun sebaliknya. Adapun tabel interpretasi koefisien korelasinya adalah sebagai berikut:

Tabel 8 Tabel Interpretasi Koefisien Pearson Correlation

No	Nilai r	Interpretasi
1	0.00 - 1.199	Sangat Rendah
2	0.20 - 0.399	Rendah
3	0.40 - 0.599	Sedang
4	0.60 - 0.799	Kuat
5	0.80 - 1.000	Sangat Kuat

Correlations - Orange			
Pearson correlation			
			Y
Filter ...			
1	+0.829	X5	Y
2	+0.297	X3	Y
3	-0.228	X4	Y
4	+0.169	X1	Y

GAMBAR 13 HASIL PEARSON CORRELATION PADA APLIKASI ORANGE

Dapat dilihat pada Gambar 13 merupakan hasil dari spearman correlation yang dilakukan di aplikasi Orange. Berdasarkan hasil tersebut menunjukkan bahwa

1. Variabel X1 bernilai +0.169 maka hubungan variabel X1 dengan variabel Y memiliki hubungan yang positif. Artinya hubungan antara variabel produksi *Crude Palm Oil/CPO* dengan variabel harga rata-rata minyak goreng domestik berbanding lurus atau searah dengan interpretasi sangat rendah.
2. Variabel X3 bernilai +0.297 maka hubungan variabel X3 yaitu variabel Y memiliki hubungan yang positif. Artinya hubungan antara variabel jumlah konsumsi minyak goreng domestik dengan variabel harga rata-rata minyak goreng domestik berbanding lurus atau searah dengan interpretasi rendah.
3. Variabel X4 bernilai -0.228 maka hubungan variabel X4 dengan variabel Y memiliki hubungan yang negatif. Artinya hubungan antara variabel jumlah ekspor minyak goreng dengan variabel harga rata-rata minyak goreng domestik berbanding terbalik atau tidak searah dengan interpretasi rendah.
4. Variabel X5 bernilai +0.829 maka hubungan variabel X5 dengan variabel Y memiliki hubungan yang positif. Artinya hubungan antara variabel harga *Crude Palm Oil/CPO* dunia dan variabel harga rata-rata minyak goreng domestik berbanding lurus atau searah dengan interpretasi sangat kuat.

### Persamaan Regresi Linear

Berdasarkan model regresi linear yang dibuat oleh peneliti pada aplikasi Orange maka dapat dilihat pada Table 8 menunjukkan bahwa output model regresi linear menghasilkan koefisien-koefisien variabelnya sebagai berikut:

Tabel 8 Output Model Regresi Linear

Hasil Persamaan Regresi - Orange	
name	coef
1 intercept	7.79082
2 X5	0.302284
3 X1	-0.0446982
4 X4	-0.0305268
5 X3	0.0602752

Berdasarkan hasil output model regresi linear di atas maka model regresi menghasilkan:

$$Y = \alpha + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \varepsilon$$

Menjadi:

$$Y = 7.79082 - 0.0446982(X_1) + 0.0602752(X_3) - 0.0305268(X_4) + 0.302284(X_5)$$

Keterangan:

Y = Harga rata-rata minyak goreng domestik (Variabel terikat)

a = Konstanta/Intercept

$\beta_1$  = Koefisien Produksi *Crude Palm Oil/CPO*

$\beta_3$  = Koefisien Jumlah Konsumsi Minyak Goreng Domestik

$\beta_4$  = Koefisien Jumlah Ekspor Minyak Goreng

$\beta_5$  = Koefisien Harga *Crude Palm Oil/CPO* Dunia

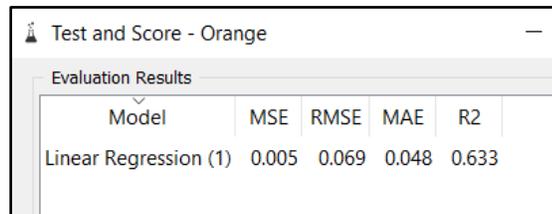
$X_1$  = Variabel Produksi *Crude Palm Oil/CPO*

$X_3$  = Variabel Jumlah Konsumsi Minyak Goreng Domestik

$X_4$  = Variabel Jumlah Ekspor Minyak Goreng

$X_5$  = Variabel Harga *Crude Palm Oil/CPO* Dunia

Tabel 9 Hasil Model Regresi Linear

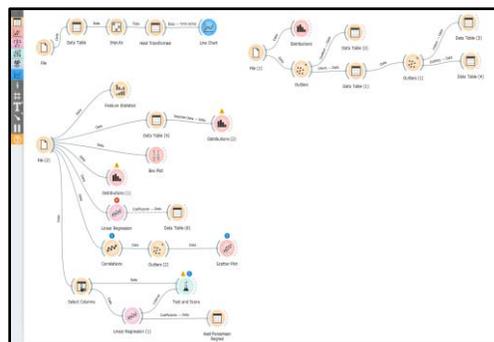


Model	MSE	RMSE	MAE	R2
Linear Regression (1)	0.005	0.069	0.048	0.633

R-square memiliki nilai 0.633 atau 63.3% artinya  $X_1$  (produksi *Crude Palm Oil/CPO*),  $X_3$  (jumlah konsumsi domestik),  $X_4$  (jumlah ekspor minyak goreng), dan  $X_5$  (harga *Crude Palm Oil/CPO* dunia) mempengaruhi 63.3% terhadap variabel Y (harga rata-rata minyak goreng domestik) sedangkan sisanya sebesar 36.7% dipengaruhi oleh faktor lain yang tidak dimasukkan ke dalam model penelitian.

## Evaluasi

Proses evaluasi dalam penelitian ini merupakan tahapan pengecekan kembali yang difokuskan kepada model atau pola yang dihasilkan dengan algoritma regresi pada aplikasi Orange apakah model sudah dapat berjalan sesuai dengan persyaratan yang ada. Dalam penelitian ini, sebelum dilakukan uji regresi linear berganda perlu melakukan uji asumsi klasik sehingga model dapat berjalan sesuai dan dapat menghasilkan output yang diperlukan. Dan pada model correlation yang ada pada aplikasi Orange, penulis menggunakan *Pearson correlation* dikarenakan data yang diolah merupakan tipe data interval atau rasio dan berdistribusi normal sehingga lebih memenuhi syarat dibandingkan opsi menggunakan *spearman correlation*. Dapat dilihat pada Gambar 14 yang merupakan hasil model data mining menggunakan algoritma regresi linear berganda yang dilakukan pada aplikasi Orange.



GAMBAR 14 HASIL MODEL DATA MINING PADA APLIKASI ORANGE

Berdasarkan hasil uji pada variabel X1 tersebut, menunjukkan bahwa ada korelasi antara produksi *CPO* dengan harga minyak goreng kemasan di Indonesia. Maka dapat disimpulkan bahwa jika produksi *CPO* naik maka harga minyak goreng kemasan di Indonesia juga naik karena hubungan korelasi nya bergerak searah. Hal tersebut dapat terjadi dikarenakan ketika harga tinggi maka produksi *CPO* akan lebih banyak karena banyaknya produsen *CPO* yang menanam sawit.

Berdasarkan hasil uji pada variabel X3, menunjukkan bahwa ada korelasi antara total konsumsi minyak goreng domestik dengan harga dari minyak goreng karena koefisien korelasinya berbanding lurus atau searah. Hal tersebut dikarenakan, ketika *demand* terus mengalami peningkatan maka harga ikut meningkat. Meskipun secara teori *supply-demand* hal ini berlawanan namun dalam beberapa kasus produk ada kemungkinan dimana kurva permintaan (*demand*) mempunyai *slope* positif dimana hal tersebut mengindikasikan bahwa hubungan antara harga dan jumlah yang diminta adalah positif, atau jika harga suatu barang naik maka jumlah yang diminta juga naik begitupun sebaliknya.

Berdasarkan hasil uji yang dilakukan pada variabel X4, hal ini menjelaskan bahwa jumlah ekspor minyak goreng naik maka harga minyak goreng domestik akan turun karena hubungan koefisiennya bernilai negatif. Hal ini berlawanan dengan teori *supply and demand*, dimana pada kasus ekspor produk minyak goreng berdasarkan data *real* yang peneliti himpun mulai dari tahun 2017 sampai dengan April 2022 serta data Kementerian Pertanian Republik Indonesia menyebutkan bahwa meskipun Indonesia merupakan negara eksportir minyak sawit utama dunia namun perkebunan kelapa sawit di Indonesia saat ini masih banyak yang dikelola oleh investor asing. Sebanyak 40% dari total lahan kelapa sawit di Indonesia dikelola oleh investor asing dimana Malaysia dan menguasai 3 juta hektar lahan kelapa sawit yang ada di Indonesia. Selain itu, hingga saat ini harga komoditas sawit dikendalikan oleh Bursa Malaysia Derivatives (BMD) dan harga minyak sawit yang di jual Indonesia juga mengacu kepada bursa komoditas yang berada di Rotterdam, Belanda. Oleh karena itu, pergerakan ekspor Indonesia masih sangat bergantung kepada investor-investor asing tersebut dan ketika harga *CPO* dunia naik maka berdampak timbulnya kekhawatiran pasokan *CPO* dari negara produsen khususnya Indonesia yang terbatas sehingga akibat dari kekhawatiran tersebut, importir membeli lebih banyak *CPO* dan Indonesia mengeksport lebih banyak yang akhirnya mendorong kenaikan harga minyak goreng domestik.

Berdasarkan hasil uji yang dilakukan pada variabel X5, hal ini mengartikan bahwa jika harga *CPO* dunia meningkat maka harga minyak goreng domestik juga ikut meningkat. Variabel X5 memiliki koefisien korelasi yang sangat kuat terhadap variabel Y dimana GAPKI menyatakan bahwa harga dari *CPO* dunia yang meningkat di dorong karena permintaan (*demand*) minyak sawit yang terus meningkat baik di dalam negeri maupun luar negeri.

## PENUTUP

### Kesimpulan

Berdasarkan analisis dan implementasi model dalam proses data mining menggunakan metode *Cross-Industry Standard Process for Data Mining (CRISP-DM)* dengan teknik analisis regresi menggunakan database faktor-faktor yang mempengaruhi harga rata-rata minyak goreng domestik sebagai berikut:

1. Hasil uji analisis regresi linear berganda dengan beberapa variabel X terhadap variabel Y menunjukkan bahwa variabel X5 yaitu harga *Crude Palm Oil/CPO* Dunia memiliki interpretasi yang sangat kuat bernilai +0.829 mempengaruhi secara positif/berbanding lurus/searah terhadap variabel Y yaitu harga rata-rata minyak goreng domestik.
2. Variabel X1 (produksi *Crude Palm Oil/CPO*), X3 (jumlah konsumsi domestik), X4 (jumlah ekspor minyak goreng), dan X5 (harga *Crude Palm Oil/CPO* dunia) mempengaruhi 63.3% terhadap variabel Y (harga rata-rata minyak goreng domestik) sedangkan sisanya sebesar 36.7% dipengaruhi oleh faktor lain yang tidak dihitung dalam penelitian.

## SARAN

1. Dalam melakukan pemodelan sebaiknya bisa ditambahkan dengan teknik data mining lainnya, tidak hanya menggunakan teknik analisis regresi sehingga dapat menghasilkan output model yang lebih optimal.
2. Untuk penelitian selanjutnya, sebaiknya lebih banyak memasukkan faktor-faktor yang mempengaruhi terhadap harga rata-rata minyak goreng domestik.

## REFERENCES

- Arhami, Muhammad dan Muhammad Nasir. 2020. *Data Mining: Algoritma dan Implementasi*. Yogyakarta: Penerbit ANDI.
- Badan Pusat Statistik. 2017. *Statistik Kelapa Sawit Indonesia 2017*. Indonesia: Badan Pusat Statistik.
- Badan Pusat Statistik. 2018. *Statistik Kelapa Sawit Indonesia 2018*. Indonesia: Badan Pusat Statistik.
- Badan Pusat Statistik. 2019. *Statistik Kelapa Sawit Indonesia 2019*. Indonesia: Badan Pusat Statistik.
- Badan Pusat Statistik. 2020. *Statistik Kelapa Sawit Indonesia 2020*. Indonesia: Badan Pusat Statistik.
- Badan Pusat Statistik. 2021. *Statistik Kelapa Sawit Indonesia 2021*. Indonesia: Badan Pusat Statistik.
- Budiharto, Widodo. 2018. *Big Data for Beginners*. Jakarta: Binus University.
- Fayyad, U.M. 1996. *Advances in Knowledge Discovery and Data Mining*. MIT Press.
- GAPKI. 2017. *Refleksi Industri Sawit 2017 & Prospek 2018*. diakses dari <https://gapki.id>, diakses pada April 2022 pada jam 12.01 WIB.
- GAPKI. 2018. *Refleksi Industri Sawit 2018 & Prospek 2019*. diakses dari <https://gapki.id>, diakses pada April 2022 pada jam 12.25 WIB.
- GAPKI. 2019. *Refleksi Industri Sawit 2019 & Prospek 2020*. diakses dari <https://gapki.id>, diakses pada April 2022 pada jam 13.31 WIB.
- GAPKI. 2020. *Refleksi Industri Sawit 2020 & Prospek 2021*. diakses dari <https://gapki.id>, diakses pada Mei 2022 pada jam 13.55 WIB.
- GAPKI. 2021. *Kinerja Industri Sawit 2021 & Prospek 2022*. diakses dari <https://gapki.id>, diakses pada Mei 2022 pada jam 14.30 WIB.
- Hariyadi, Purwiyatno. 2015. *Mengenal Minyak Sawit dengan Beberapa Karakter Unggulnya*. Jakarta: GAPKI.
- Indonesia. 2021. *Peraturan Menteri Perdagangan Republik Indonesia nomor 19 tahun 2021 tentang Kebijakan dan Pengaturan Ekspor*. Jakarta.
- Inmon, William H. 2005. *Building Data Warehouse 3th Edition*. Canada: John Wiley & Sons.
- Ketaren, S. 2008. *Pengantar Teknologi Minyak dan Lemak Pangan*. Jakarta: UI-Press.
- Ketaren, S. 2012. *Pengantar Teknologi Minyak dan Lemak Pangan*. Jakarta: UI-Press.
- Kusrini dan Emha T.L., 2009. *Algoritma Data Mining*. Yogyakarta: Penerbit ANDI.
- Larose, Daniel T. 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*. John Willey & Sons, Inc.
- Manuhutu, Abraham. 2021. *Data Mining dan Aplikasinya*. Bandung: Widina Bhakti Persada.
- Pahan, Iyung. 2008. *Panduan Lengkap Kelapa Sawit*. Jakarta: Niaga Swadaya.
- PIHPS Nasional. 2022. *Laporan Monthly Harga Minyak Goreng Kemasan Kg periode 2017 sampai dengan 2022*. diakses dari <https://hargapangan.id>, diakses pada Juli 2022 pada jam 15.10 WIB
- Provost, Foster and Tom Fawcett. 2013. *Data Science for Business*. California: O'Reilly Media.
- Rambe, Khoiru Rizqy dan Nunung Kusnandi. 2018 *Permintaan dan Penawaran Minyak Goreng Sawit Indonesia*. Departemen Agribisnis Institut Pertanian Bogor, Vol. 8, No. 1.

- Risti, I. 2016. Uji Kualitas Minyak Goreng Curah Dan Minyak Goreng Kemasan di Manado. *Jurnal Ilmiah Farmasi-UNSRAT*, 5(4).
- Siregar MA, Sembiring SA, Ramli. 2014. *The Price of Palm-Cooking Oil in Indonesia: Antecedents and Consequences on the International Price and the Export Volume of CPO. Journal of Economics and Sustainable Development*. 5(23): 227-234.
- Turban, dkk. 2005. *Decision Support Systems and Intelligent Systems*. Yogyakarta: Penerbit ANDI.
- Turban, Efraim & Linda Volonino. 2010. *Information Technology for Management edisi Ketujuh*. Asia: John Wiley & Sons.
- World Bank Commodity Price*. 2022. *World Bank Commodity Price Data (The Pink Sheet) CPO Price*. diakses dari <https://www.worldbank.org>, diakses pada Agustus 2022 pada jam 20.00 WIB.

