

Implementasi Regresi Linear Untuk Memprediksi Hasil Impor Jumlah Barang Konsumsi Tahun 2021-2036

Anthonius¹, Charles Calvin King Luise^{2*}, Juven Prisselix³

^{1,2,3} Teknik Perangkat Lunak, Universitas Universal

* Corresponding author E-mail: charleskingluise@uvers.ac.id

Article Info

Article history:

Received 30-11-2021

Revised 06-12-2021

Accepted 06-12-2021

Keyword:

CRISP-DM,
Data Mining,
Impor,
Prediksi.

ABSTRACT

Import is a very important activity for a country and the value always changing every year, this thing can affect the foreign exchange of a country. When a country consumed a lot goods, it resulted in the country to do a lot of import activity as well. By doing so, the economy would be unbalanced and so the country will have a lot of debts. With that in mind, predicting the import activity in the future can be essential to prevent the country from being unbalanced in terms of economy. The purpose of this research is to predict the future import activity by using Linear Regression Algorithm and applying Cross-Industry Standard Process for Data Mining (CRISP-DM) method. This method uses to find the value of regression equation, as well as getting error analysis through accuracy of predictions using MAD, MSE, and MAPE using R. Studio software. The results that was obtained for 2021 amounted to 6745.298 tons and 2022 amounted to 6578.703 tons. As for the resulting error analysis, MAD value is 2652.901, MSE value is 10002316 and MAPE value is 255.17%.



Copyright © 2021. This is an open access article under the [CC BY](https://creativecommons.org/licenses/by/4.0/) license.

I. PENDAHULUAN

Impor bila merujuk pada Peraturan Pemerintah Nomor 10 Tahun 2021 memiliki arti yaitu kegiatan memasukkan barang ke dalam daerah pabean. Dengan kata lain impor berarti kegiatan yang memasukkan barang atau jasa dari luar negeri untuk diedarkan ke dalam negeri. Pembelian impor juga pada umumnya adalah pembelian barang-barang yang tidak bisa diproduksi dalam negeri [1].

Tidak ada negara yang bisa berdiri sendiri tanpa adanya melakukan hubungan perdagangan dengan negara lain, Indonesia sebagai negara yang masih berkembang juga tidak terkecuali. Namun, bukan berarti jumlah barang yang diimpor akan terus bertambah dan tidak akan berkurang. Data dari BPS menunjukkan bahwa kegiatan impor Indonesia mengalami kenaikan dari tahun sebelumnya pada April 2021 mengalami kenaikan disbanding dengan tahun sebelumnya pada bulan yang sama sebesar 22,1% untuk impor non migas [3]. Penelitian ini bertujuan untuk memprediksi apakah kedepannya kegiatan impor Indonesia akan berkurang setelah sebelumnya mengalami kenaikan yang cukup signifikan dari tahun sebelumnya.

Penelitian mengenai prediksi impor sebelumnya sudah pernah dilakukan oleh "Gheybla Najla Ayuni, Devi Fitriannah" dengan judul "Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ"[6], kemudian penelitian yang dilakukan oleh "Nova Kusumawati, Fitri Marisa, Indra Dharma Wijaya" dengan judul penelitian "Prediksi Kurs Rupiah Terhadap Dollar Amerika Dengan Menggunakan Metode Regresi Linear"[5], dan penelitian terakhir lainnya yang dilakukan oleh "Ria Iswandari, Lola Anandya Inke, Indira Hapsari" dengan judul penelitian "Analisis Peramalan Produksi Singkong dan Kelayakan Finansial Agroindustri Mocaf Di Provinsi Lampung"[7], maka dari itu kami melakukan prediksi impor barang konsumsi ke Indonesia pada tahun mendatang yaitu 2021-2036.

Manfaat dilakukannya "Implementasi Regresi Linear Untuk Memprediksi Hasil Impor Jumlah Barang Konsumsi Tahun 2021-2036" adalah sebagai berikut:

1. Untuk mencari prediksi jumlah barang impor pada tahun 2021-2036. Bila penelitian berhasil, maka prediksi untuk hasil impor dapat dijadikan referensi untuk prediksi hasil impor tahun-tahun mendatang tidak hanya pada tahun 2021.

2. Mengetahui produktivitas dalam negeri di masa mendatang. Impor bila mengalami penurunan kemungkinan produktivitas dalam negeri mengalami kenaikan dari segi kualitas dan juga kemungkinan lebih mampu memenuhi kebutuhan dalam negeri.

3. Memprediksi kondisi negara. Pandemi yang sudah melanda dunia semenjak akhir tahun 2019 membuat banyak negara mengalami kerugian dari segi ekonomi karena produktivitas juga menurun. Data ini bisa dijadikan acuan bila negara mengalami kenaikan yang signifikan dari tahun sebelumnya, maka terdapat kondisi tertentu yang memaksa negara untuk lebih banyak melakukan kegiatan impor guna memenuhi kebutuhannya, sebaliknya bila negara mengalami penurunan dalam kegiatan impor bisa berarti bahwa keadaan sudah kembali normal.

Indonesia adalah negara berkembang yang memiliki banyak sumber daya melimpah. Namun, dengan banyak sumber daya yang melimpah tersebut kegiatan impor tetap banyak dilakukan. Hal itu karena biaya produksi yang mahal dan sebagai negara berkembang, bila memilih antara membeli barang yang sudah jadi namun lebih murah dibanding produksi sendiri maka negara akan memilih membeli. Terdapat juga barang yang tidak tersedia dalam negeri sehingga hal tersebut memaksa negara untuk melakukan kegiatan impor [4].

1. Bagaimana prediksi hasil impor dapat mempengaruhi hasil kegiatan impor?
2. Apakah mungkin untuk menurunkan kegiatan impor dengan prediksi hasil impor?

Dataset yang digunakan berdasarkan data dari Badan Pusat Statistik yang diupdate pada tanggal 21 Juli 2021. Hasil yang didapat adalah prediksi hingga tahun 2030 didapat dengan menggunakan Teknik Data Mining Regresi Linear menggunakan dua variabel. *Variable Independent* (X) dalam hal ini adalah Jumlah Barang Konsumsi dan Tahun sebagai *variabel dependent* (Y). Tujuan penelitian ini dilakukan dengan harapan agar pemerintah dapat lebih mengembangkan kualitas produk dalam negeri. Dengan berkembangnya kualitas dalam negeri, maka kegiatan impor juga dapat dikurangi dan lebih menyeimbangkan keuangan negara.

II. METODE

Tahapan yang di lakukan di metode ini menggunakan *Cross Industry Standar Process for Data Mining (CRISP-DM)* dan berikut langkah langkahnya:

A. Business Understanding

Pada tahapan pemahaman bisnis ada beberapa hal yang dilakukan antara lain, menentukan tujuan dan persyaratan dengan jelas secara keseluruhan, dan selanjutnya mempersiapkan strategi awal untuk mencapai tujuan tersebut.

B. Data Understanding

Pada tahapan pemahaman data, ada beberapa hal yang dilakukan antara lain, memeriksa dan mengumpulkan data, sehingga dapat mendeskripsikan dan memilih data

C. Data Preparation

Setelah data dikumpulkan, data diidentifikasi di pilah dan kemudian di bagi sesuai dengan yang di perlukan.

D. Modeling

Setelah data siap, tahap selanjutnya adalah melakukan modeling menggunakan aplikasi algoritma atau R. Studio untuk mencari, mengidentifikasi, dan menampilkan pola. Pemilihan algoritma berdasarkan pada tipe data, algoritma yang di gunakan adalah prediksi, melihat hubungan asosiatif. Adapun dalam Penelitian ini menggunakan algoritma regresi linear berganda yang bertujuan untuk melakukan prediksi.

E. Evaluation

Melakukan evaluasi akurasi prediksi menggunakan metode *Mean Absolute Deviation (MAD)*, *Mean Squared Error (MSE)*, dan *Mean Absolute Percent Error (MAPE)*. Hasil evaluasi dapat digunakan untuk mendapatkan kesimpulan dari kualitas hasil prediksi. Secara sistematis MAD dirumuskan seperti pada Persamaan 1:

$$MAD = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (1)$$

y = nilai data aktual

\hat{y}_i = nilai data prediksi

n = jumlah data

Dari rumus ini selisih nilai data aktual dengan data prediksi atau data peramalan. Secara sistematis MSE dirumuskan seperti pada Persamaan 2:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

y_i = nilai data Aktual

\hat{y}_i = nilai data peramalan

n = banyaknya data

Pengurangan nilai data aktual dengan data peramalan dan hasilnya dikuadratkan kemudian dijumlahkan secara keseluruhan. Secara sistematis MAPE dirumuskan seperti pada Persamaan 3:

$$MAPE = \frac{1}{n} \sum_{t=1}^n A \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100 \% \quad (3)$$

n = ukuran sampel

y_i = nilai data aktual

\hat{y}_i = nilai data peramalan

Dari rumus ini, selisih data aktual dengan data peramalan dibagi dengan data aktual, setelah itu, hasilnya di kali kan dengan 100% kemudian nilainya di mutlakkan (absolut). Artinya MAPE akan selalu bernilai positif.

III. HASIL DAN PEMBAHASAN

A. Business Understanding

Hasil Prediksi Impor di Indonesia pada tahun 2022 di dapat melalui *variable independent* (bebas) dan *variable dependen*(terikat). Tujuan dari prediksi berikut ini menggunakan aplikasi R. Studio merupakan keuntungan buat

Indonesia karena negara bisa meminimalisir kerugian berupa kelebihan barang impor atau kekurangan barang impor.

Strategi nya dengan melakukan perhitungan / prediksi Impor di Indonesia dan mengetahui hasilnya menggunakan metode MAD, MSE dan MAPE.

B. Data Understanding

Data Impor di Indonesia di dapatkan melalui website Badan Pusat Statistik [2], berikut atributnya antara lain:

Tabel 1. Data Impor Dengan Atribut Tahun dan Jumlah Barang Konsumsi

Tahun	Jumlah Barang Konsumsi	Tahun	Jumlah Barang Konsumsi
1989	706,1	2005	5 562,1
1990	584,0	2006	4 706,6
1991	734,8	2007	6 714,4
1992	1 255,2	2008	5 368,1
1993	799,5	2009	4 056,6
1994	1 899,8	2010	5 604,5
1995	3 396,3	2011	8 110,5
1996	4 322,0	2012	6 966,7
1997	2 338,3	2013	5 285,7
1998	4 158,6	2014	5 599,1
1999	7 324,5	2015	4 929,2
2000	5 241,2	2016	5 899,5
2001	4 071,2	2017	5 315,0
2002	5 643,4	2018	7 398,9
2003	4 903,4	2019	5 255,6
2004	4 749,9	2020	5 205,0

Data yang di peroleh dari *Website* BPS (Badan Pusat Statistik) dari tahun 1989-2020, dengan Format Excel dan melalui pemilihan data, maka di ambil 2 atribut saja, yaitu: Tahun dan Jumlah Barang Konsumsi.

Hasil korelasi tahun dan JBK menggunakan fungsi *cor* adalah 0.7492483 dan dari hasil tersebut diketahui bahwa hasil tersebut mendekati angka 1 maka tingkat pengaruh yang dimiliki variabel *x* dan *y* tinggi.

C. Data Preparation

Langkah-langkah yang di gunakan pada tahap data preparation ini antara lain:

a. Select Data

Tahapan ini adalah tahapan memilih data yang di perlukan untuk di perhitungkan atau di prediksi, gambar ada pada Tabel 1.

b. Format Data

Tahap penentuan Atribut dalam data yang akan di gunakan atau di eksekusi di program R. Studio ada di tabel 1.

c. Data Separation

Tahap ini merupakan taha pembagian data menjadi dua bagian, yaitu data latih dan data uji. Data Latih yang digunakan dari Data Tahun 1989 – 2010.

Tabel 3. Data Latih

Tahun	Jumlah Barang Konsumsi
1989	706,1
1990	584,0

1991	734,8
1992	1 255,2
1993	799,5
1994	1 899,8
1995	3 396,3
1996	4 322,0
1997	2 338,3
1998	4 158,6
1999	7 324,5
2000	5 241,2
2001	4 071,2
2002	5 643,4
2003	4 903,4
2004	4 749,9

Data Uji yang kita gunakan dari Tahun 2011-2020

Tabel 4. Data Uji

Tahun	Jumlah Barang Konsumsi
2005	5 562,1
2006	4 706,6
2007	6 714,4
2008	5 368,1
2009	4 056,6
2010	5 604,5
2011	8 110,5
2012	6 966,7
2013	5 285,7
2014	5 599,1
2015	4 929,2
2016	5 899,5
2017	5 315,0
2018	7 398,9
2019	5 255,6
2020	5 205,0

D. Modeling

Dalam melakukan prediksi ke Tahun depan, terdeteksi bahwa di dapatkan hasil:

Tabel 4. Hasil Prediksi

Tahun	Hasil Prediksi
2021	449.1919
2022	823.7013
2023	1198.2107
2024	1572.7201
2025	1947.2296
2026	2321.7390
2027	2696.2484
2028	3070.7578
2029	3445.2672
2030	3819.7766
2031	4194.2860
2032	4568.7954
2033	4943.3049
2034	5317.8143
2035	5692.3237
2036	6066.8331

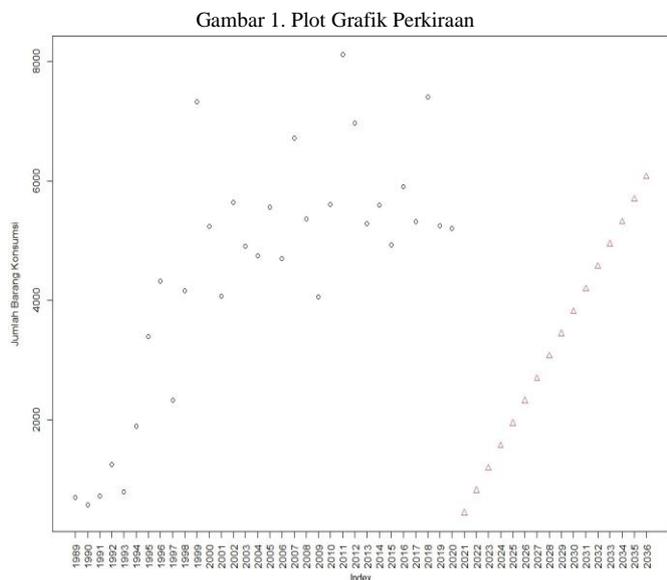
Model Prediksi:

```
modelfit <- lm (datauji ~ tahun)
```

Model yang dibentuk menggunakan fungsi *lm* (*Linear Models*). Variabel yang digunakan menggunakan datauji di Tabel 4. Selanjutnya di buat pembentukan fungsi *predict*:

```
datanew <- predict(modelfit, newdata=databaru)
```

Kemudian keluarlah hasil dari plot prediksi tersebut:



Perkiraan pada Gambar 1. terdapat 2 bentuk dan warna berbeda titik bulat dan segitiga merah, dan maksud dari bentuk tersebut juga beda beda. Titik hitam bulat menggambarkan data aktual (Tahun 2005-2020) dan kemudian segitiga merah menggambarkan data prediksi (Tahun 2021-2036) dengan menggunakan R. Studio. Dan mendapatkan hasil pada tahun 2021 sebanyak 449.1919 Ton, pada tahun 2022 sebanyak 823.7013 ton, pada tahun 2023 sebanyak 1198.2107 ton, pada tahun 2024 sebanyak 1572.7201 ton, pada tahun 2025 sebanyak 1947.2296, pada tahun 2026 sebanyak 2321.7390, pada tahun 2027 sebanyak 2696.2484 ton, pada tahun 2028 sebanyak 3070.7578 ton, pada tahun 2029 sebanyak 3445.2672 ton, pada tahun 2030 sebanyak 3819.7766 ton, pada tahun 2031 sebanyak 4194.2860, pada tahun 2032 sebanyak 4568.7954 ton, pada tahun 2033 sebanyak 4943.3049 ton, pada tahun 2034 sebanyak 5317.8143 ton, pada tahun 2035 sebanyak 5692.3237 ton, pada tahun 2036 sebanyak 6066.8331 ton.

E. Evaluation

Tahap Penutupan ini biasanya untuk melakukan atau mengetahui keakuratan dari prediksi suatu program, akan tetapi dalam prediksi tidak ada yang akurat, karena terjadi ketidaksesuaian data aktual dengan data hasil dari prediksi.[4]. Karena itu dilakukan perhitungan kesalahan atau error yang menggunakan data selisih dan menggunakan metode analisis kesalahan *Mean Absolute Deviation* (MAD),

Mean Squared Error (MSE), dan *Mean Absolute Percent Error* (MAPE).

a. MAD (*Mean Absolute Deviation*), adalah rata-rata kesalahan mutlak selama periode tertentu tanpa memperhatikan apakah hasil prediksi lebih besar atau lebih kecil dibanding aktualnya.

b. MSE (*Mean Square Error*), adalah rata-rata kesalahan kuadrat yang dihitung dengan semua kesalahan peramalan pada setiap periode.

c. MAPE (*Mean Absolute Percentage Error*), adalah rata-rata presentase kesalahan mutlak.

Sebelum melakukan langkah prediksi, terlebih dahulu akan membuat model prediksi pada R. Studio, dan bisa di lihat modelnya berikut ini.

Model Prediksi:

```
modelfit <- lm(datalatih [,2] ~ datalatih [,1])
```

Setelah pembuatan model prediksi, maka langsung di lakukan prediksi di R. Studio di bawah ini.

Predict:

```
jbkpred <- predict(modelfit, newdata=datauji [,1])
```

Setelah melakukan prediksi dan mendapatkan hasilnya maka hasilnya terdapat di tabel 4. Kemudian di lakukan perhitungan selisih yang bertujuan mengetahui selisih kesalahan (*error*) dari prediksi yang di lakukan. Berikut hasilnya.

Tabel Selisih:

Tabel 5. Hasil Selisih

Tahun	Selisih
2021	5112.9081
2022	3882.8987
2023	5516.1893
2024	3795.3799
2025	2109.3704
2026	3282.7610
2027	5414.2516
2028	3895.9422
2029	1840.4328
2030	1779.3234
2031	734.9140
2032	1330.7046
2033	371.6951
2034	2081.0857
2035	436.7237
2036	861.8331

Selisih:

$$selisih = abs(datanew-datauji)$$

dengan selisih yang ada, databaru di kurangi dengan datauji. Hasil data selisih digunakan untuk mencari seberapa besar selisih error pada data prediksi yang didapat dari hitungan error MAD, MSE, dan MAPE.maka terciptalah hasil MAD dengan jumlah 2652.901, kemudian MSE dengan

jumlah 10002316, dan yang terakhir perhitungan error MAPE dengan jumlah 255.1756. Nilai Aktualnya sebesar 2652.901 yang di peroleh dari rata rata kesalahan. Nilai rata rata kesalahan berpangkat atau MSE sebesar 10002316. Nilai MAPE atau nilai yang menunjukkan rata rata kesalahan dalam prediksi mencapai 255.17%. Data MAPE ini menunjukkan bahwa error dari nilai prediksi yang di lakukan ini akan membuat error sebanyak 255.17% atau bisa di sebut juga dengan tingkat ketidakakuratan nya mencapai 255.17%.

IV. KESIMPULAN

Hasil prediksi barang impor di Indonesia menggunakan R. Studio dengan Teknik Regresi Linear, mendapatkan hasil perkiraan dengan, pada tahun 2021 sebanyak 449.1919 Ton, pada tahun 2022 sebanyak 823.7013 ton, pada tahun 2023 sebanyak 1198.2107 ton, pada tahun 2024 sebanyak 1572.7201 ton, pada tahun 2025 sebanyak 1947.2296 ton, pada tahun 2026 sebanyak 2321.7390 ton, pada tahun 2027 sebanyak 2696.2484 ton, pada tahun 2028 sebanyak 3070.7578 ton, pada tahun 2029 sebanyak 3445.2672 ton, pada tahun 2030 sebanyak 3819.7766 ton, pada tahun 2031 sebanyak 4194.2860 ton, pada tahun 2032 sebanyak 4568.7954 ton, pada tahun 2033 sebanyak 4943.3049 ton, pada tahun 2034 sebanyak 5317.8143 ton, pada tahun 2035 sebanyak 5692.3237 ton, pada tahun 2036 sebanyak 6066.8331 ton.

Hasil Prediksi ini dari tahun ketahun di perkirakan akan semakin meningkat dan meningkat, dan di harapkan agar

pemerintah Indonesia bisa mengatur tingkat impor barang ke Indonesia, karena jika impor yang berlebihan bisa berdampak yang sangat buruk bagi negara.

DAFTAR PUSTAKA

- [1] M. Idris, "Apa Itu Impor: Pengertian, Tujuan, dan Contohnya", Kompas, April 7, 2021, [online]. Available: <https://money.kompas.com/read/2021/04/07/080257426/apa-ituimporpengertian-tujuan-dan-contohnya?page=all>. [Accessed Nov 26, 2021].
- [2] D. Sensus, "Impor Barang Konsumsi, 1989-2020", Badan Pusat Statistik [Accessed Nov 28, 2021]. Available: <https://www.bps.go.id/statictable/2014/09/08/1040/impor-barang-konsumsi-1989-2020.html>.
- [3] C. M. Annur, "Negara Asal Produk Impor Non-Migas Terbesar RI Berasal dari Tiongkok", Databoks, May 24, 2021, [online]. Available: <https://databoks.katadata.co.id/datapublish/2021/05/24/negara-asal-produk-impor-non-migas-terbesar-ri-berasal-dari-tiongkok>. [Accessed Nov 26, 2021].
- [4] M. Sitepu, "Mungkinkah satu negara tak melakukan impor pangan atau produk lain?", BBC News Indonesia, Nov 7, 2018, [online]. Available: <https://www.bbc.com/indonesia/indonesia46110555>. [Accessed Nov 26, 2021].
- [5] N. Kusumawati, F. Marisa, I. D. Wijaya, "Prediksi Kurs Rupiah Terhadap Dollar Amerika Dengan Menggunakan Metode Regresi Linear" Vol 2 No 3 Desember 2017.
- [6] G. N. Ayuni, and D. Fitriah, "Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ" vol. 14 no. 2, 2019.
- [7] R. Iswandari, L. A. Inke, and I. Hapsari "ANALISIS PERAMALAN PRODUKSI SINGKONG DAN KELAYAKAN FINANSIAL AGROINDUSTRI MOCAF DI PROVINSI LAMPUNG" Vol. 14 No. 1 Februari 2021.