

Minimalisasi Resiko Kualitas Barang dengan Metode Algoritma Regresi Linier dalam Memprediksi Barang *Defect*

U. Darmanto Soer^{1*)}, Sifa Fauziah²⁾, Uswatun Hasanah³⁾

¹⁾²⁾³⁾Program Studi Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa

^{*)}Correspondence author: darmantosoer@pelitabangsa.ac.id, Jakarta, Indonesia

DOI: <https://doi.org/10.37012/jtik.v10i1.2273>

Abstrak

Barang merupakan sesuatu yang dapat ditawarkan ke pelanggan untuk mendapatkan perhatian untuk dibeli atau digunakan agar dapat memenuhi keinginan dan kebutuhan pelanggan. Pemanfaatan data produk dan *defect* dapat digunakan dalam melakukan proses tahapan *data mining* dan pemodelan untuk memprediksi jumlah produk *defect* di suatu waktu. Penerapan model persamaan algoritma Regresi Linear sederhana dapat diimplementasikan dimana hasilnya juga menunjukkan sebuah wawasan baru bagi kebutuhan prediksi terhadap jumlah produk *defect*. Model persamaan Regresi Linear sederhana setelah dibandingkan hasil perhitungan secara aktual (observasi) dan juga dengan aplikasi Rapid Miner secara umum menunjukkan kemiripan hasil. Evaluasi dan pengujian performa menggunakan aplikasi Rapid Miner juga dapat menghasilkan gambaran yang relevan dengan skenario yang dimodelkan. Nilai (*Root Mean Squared*) RMSE juga didapat saat melakukan evaluasi performa model yang diterapkan, dengan nilai RMSE sebesar 0,483 dengan standar deviasi $\pm 0,0$. Rekomendasi yang diberikan adalah: Perbanyak data dalam rentang bulan lebih panjang yang akan digunakan dalam penelitian selanjutnya agar dapat meningkatkan nilai akurasi dari model prediksi terhadap jumlah produk *defect* dimasa mendatang. Pengembangan model dengan menggabungkan beberapa metode dan algoritma yang dapat digunakan dalam memprediksi sesuatu sehingga diharapkan memberikan hasil yang lebih variatif dan bisa dimanfaatkan dalam pengambilan kebijakan dan strategi pengendalian jumlah produk *defect* dan meningkatkan nilai RSME dari model prediksi yang diharapkan.

Kata Kunci: Kualitas Barang, Regresi Linier, Barang *Defect*, *Data Mining*, Rapid Miner

Abstract

Goods are something that can be offered to customers to get attention to be purchased or used in order to meet customer desires and needs. Utilization of product and defect data can be used in carrying out the data mining and modeling stages to predict the number of defective products at a time. The application of a simple Linear Regression algorithm equation model can be implemented where the results also show new insights for the need to predict the number of defective products. The simple Linear Regression equation model after comparing the actual calculation results (observations) and also with the Rapid Miner application generally shows similar results. Evaluation and performance testing using the Rapid Miner application can also produce a picture that is relevant to the modeled scenario. The RMSE (Root Mean Squared) value was also obtained when evaluating the performance of the applied model, with an RMSE value of 0.483 with a standard deviation of ± 0.0 . The recommendations given are: Increase data in a longer month range that will be used in further research in order to increase the accuracy value of the prediction model for the number of defective products in the future. Model development by combining several methods and algorithms that can be used in predicting something so that it is expected to provide more varied results and can be used in making policies and strategies to control the number of defective products and increase the RSME value of the expected prediction model.

Keywords: *Quality of Goods, Linear Regression, Defective Goods, Data Mining, Rapid Miner*

PENDAHULUAN

Barang merupakan sesuatu yang dapat ditawarkan ke pelanggan untuk mendapatkan perhatian untuk dibeli atau digunakan agar dapat memenuhi keinginan dan kebutuhan pelanggan. Dalam strategi peningkatan barang dan pengendalian kualitas produk langkah yang pertama dilakukan adalah penilaian kualitas barang, hal ini penting karena barang akan dijual ke pelanggan. Pada strategi produk yang perlu diingat yang berkaitan dengan produk secara utuh, mulai dari nama produk, bentuk, isi, dan pembungkus. Dalam arti yang sederhana produk dikatakan sesuatu yang dapat memenuhi kebutuhan dan keinginan konsumen.

Dalam pengelolaan barang, kemungkinan dapat terjadi penumpukan stok barang dan kekurangan stok barang di produk yang lain. Maka dari itu dibutuhkan suatu analisis yang dapat meramalkan kebutuhan barang untuk periode mendatang. Banyaknya penumpukan barang terjadi setiap hari membuat data produk ini dapat diolah menjadi suatu informasi dengan menggunakan metode yang tepat, dengan begitu perusahaan dapat menentukan strategi pengendalian kualitas. Akan tetapi masih banyak pengelola yang belum memanfaatkan secara maksimal data dari produk yang ada. Data dibiarkan menumpuk begitu saja tanpa dilakukan analisa lebih lanjut sehingga berpotensi pihak pengelola kurang mengetahui hal esensial terhadap terjadinya barang *defect*. Berdasarkan beberapa penjelasan di atas menunjukkan betapa pentingnya kepuasan pelanggan untuk kelangsungan bisnis di masa yang akan datang. Pelanggan memiliki peran yang sangat penting maka ketika ada klaim atau *trouble report* dari pelanggan harus secepatnya di respon dan melakukan perbaikan agar *trouble report* dan klaim tidak terulang kembali.

Berdasarkan dari uraian di atas maka pada penelitian ini, akan dikembangkan system yang dapat mempermudah perusahaan dalam hal memprediksikan strategi apa saja yang tepat untuk memasarkan produk. Estimasi adalah suatu metode untuk memperkirakan nilai karakteristik populasi dengan memakai nilai sampel dan yang dapat memodelkan persamaan untuk menghitung estimasi yakni algoritma *linear regression* (Regresi linier). Algoritma tersebut digunakan untuk memodelkan hubungan antara dua variabel dengan memasang persamaan linear untuk mengamati data. *Linear Regression* berusaha menemukan hubungan

matematis antar variabel. Didasari oleh uraian diatas maka peneliti perlu melakukan penelitian dengan judul “Minimalisasi Resiko Kualitas Barang dengan Metode Algoritma Regresi Linier dalam Memprediksi Barang *Defect*”

METODE

Dalam melakukan penelitian ini tahapan yang dilakukan sebagai berikut:

1. Pengumpulan Data

Mengumpulkan data yang dibutuhkan guna membangun untuk mendukung keputusan penentuann kebutuhan stock barang, serta mengidentifikasi masalah kualitas datanya untuk mendapatkan bagian yang menarik dari data yang dapat digunakan dalam penelitian untuk mendapatkan informasi yang tersembunyi. Teknik yang digunakan dalam penelitian ini adalah:

a. Observasi

Peneliti melakukan pengamatan dan mencatat hal-hal yang berkaitan dengan bahan penelitian. Pengamatan dilakukan dengan datang langsung dan melihat bagaimana proses pengambilan data.

b. Studi Pustaka

Dalam penelitian ini peneliti mempelajari sumber informasi yang diambil dari buku referensi dan jurnal ilmiah yang berkaitan dengan penelitian ini terutama yaitu *data mining*, estimasi, dan *regrensi linier*.

2. Pengolahan Data

Tahap ini merupakan tahap pengelolaan data berkaitan dengan pencarian subyek penelitian, dengan melakukan pemilihan tabel, record, dan atribut-atribut data, termasuk juga proses *cleaning*, *selection* dan transformasi data yang kemudian akan dijadikan masukan pada tahap *data mining*. Pengolahan data ini menggunakan metode *Association rule estimasi* .

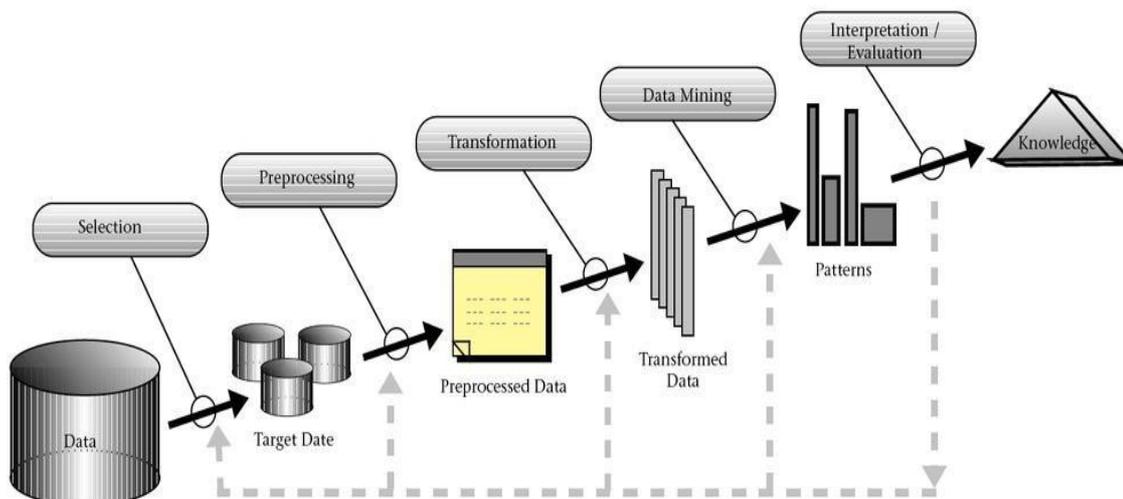
3. Pemodelan

Pemodelan pada penelitian ini menggunakan *data mining*. Teknik *estimasi* menggunakan algoritma regrensi linier. Teknik ini dipilih karena merupakan metode

yang umum dipakai pada penelitian *data mining* untuk mencari seleksi aturan, maupun sebuah prediksi

4. Evaluasi Dan Hasil

Pada tahap ini merupakan evaluasi dan hasil penelitian yang telah dilakukan sehingga dapat membantu perusahaan dalam pengambilan keputusan. Pada tahap ini pengujian dilakukan dengan Rapid Miner. Tahapan penelitian ini digambarkan pada Gambar 1.



Gambar 1. Tahapan Pemrosesan Data

Data Mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data dengan melakukan penggalian pola-pola dari data dengan tujuan untuk memanipulasi data dengan informasi yang lebih berharga yang diperoleh dengan cara mengekstraksi dari data yang ada dalam basisdata (Wibisono, Rofik, & Purwanto, 2019). *Data Mining* dianggap penting pada industri informasi karena tersedia jumlah informasi besar. Ada kebutuhan yang semakin besar untuk mengubah data tersebut menjadi informasi dan pengetahuan yang berguna (Novy, 2018). *Data Mining* adalah analisis dari data yang berjumlah besar atau kompleks dengan tujuan untuk menemukan pola atau kecenderungan yang penting yang biasanya tidak disadari keberadaannya (Ramadhan & Safitri, 2019). Dari pengertian di atas dapat disimpulkan bahwa *Data Mining* adalah proses pembelajaran komputer (*maching*

learning) untuk mengekstraksi pengetahuan (*Knowledge*) dan menemukan pola secara otomatis dari berbagai database besar.

Regresi Linier digunakan untuk mengetahui bagaimana variabel dependen/kriteria dapat diprediksi melalui variabel independen atau variabel prediktor, secara individual (Herwanto, Widiyaningtyas, & Indriana, 2019). Dampak dari penggunaan analisis regresi dapat digunakan untuk memutuskan apakah naik dan menurunnya variabel dependen dapat mempengaruhi kenaikan dan penurunan variabel independent. Meningkatkan keadaan variabel dependen dapat dilakukan dengan meningkatkan variabel independent dan sebaliknya. Metode Kuadrat terkecil (*least square method*) adalah metode yang paling populer untuk menetapkan persamaan regresi linier sederhana.

HASIL DAN PEMBAHASAN

Tahapan awal yang dilakukan pada penelitian ini adalah mempersiapkan data yang kemudian dilakukan tahapan pra-pemrosesan data seperti *data cleaning*, *data selection*, dan *data transformation* pada data produksi. Total sebanyak **1500** baris data yang dilakukan *cleaning* sehingga untuk proses berikutnya hanya tinggal tersisa data yang sudah bersih dengan total yakni **1000** baris data. Kemudian dilakukan agregasi data per bulan sehingga dihasilkan data yang akan dimodelkan seperti pada tabel 1 berikut.

Tabel 1. Agregasi Dataset Perbulan

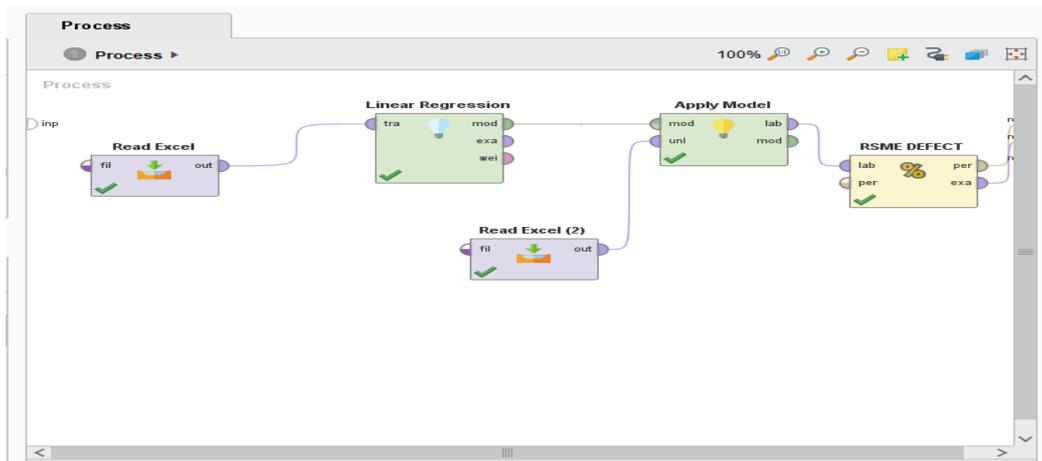
No	QTY Defect (Variabel X)	QTY Produk (Variabel Y)
1	6	12075797
2	1	10618811
3	8	12149821
4	1	13940285
5	8	12467690
6	5	12168092

Untuk selanjutnya hasil dari agregasi data tersebut akan dijadikan data latih (*training*) pada tahap pemodelan menggunakan algoritma Regresi Linier sederhana. Tahapan evaluasi terhadap data uji (*testing*) yang digunakan untuk menilai model persamaan algoritma Regresi Linier sederhana yang diterapkan dalam pembentukan prediksi produk *defect* dari suatu objek baru dengan keakurasian yang tepat.

Tabel 2. Data Training untuk Pemodelan Regresi Linier

QTY Defect (Variabel X)	QTY Produk (Variabel Y)
6	12075797
1	10618811
8	12149821
1	13940285
8	12467690
5	12168092

Penelitian ini memanfaatkan aplikasi Rapid Miner. Hasil pengujian yang dilakukan menggunakan aplikasi Rapid Miner adalah dengan menerapkan langkah-langkah sebagai berikut:



Gambar 2. Tahapan Rapidminer

Hasil nilai sample prediksi berdasar pada aplikasi Rapid Miner dapat dilihat pada tabel 3 sebagai berikut:

Tabel 3. Nilai Sampel Prediksi

Row No.	QTY Produk ...	prediction(Q...	QTY Defect (Variabel X)
1	12239321	12239320.977	3
2	12229502	12229501.974	10
3	12240724	12240723.692	2
4	12232308	12232307.403	8
5	12237919	12237918.262	4
6	12229502	12229501.974	10
7	12229502	12229501.974	10
8	12242127	12242126.407	1
9	12236516	12236515.548	5
10	12233711	12233710.118	7

Hasil nilai sample RMSE berdasar pada aplikasi Rapid Miner sebagai berikut:

root_mean_squared_error

root_mean_squared_error: 0.483 +/- 0.000

1. Melakukan import data yang diperlukan untuk proses pada tools Rapid Miner. Pada aplikasi Rapid Miner pilih dan klik **Import Data**, kemudian pilih data yang akan dipakai serta menentukan attribut dan label yang akan digunakan.
2. Klik menu **Design**, pada tampilan proses, tambahkan dataset latih (*training*) dan dataset uji (*testing*) pada folder ke layar tampilan proses.
3. Berikutnya pada menu **Modelling**, dalam submenu **Predictive**, pilih operator fungsi *Linear Regression*, untuk menerapkan persamaan algoritma Regresi Linear sederhana terhadap proses prediksi objek yang akan dilakukan.
4. Selanjutnya pada menu **Scoring** pilih operator *Apply Model* dan drag ke layar tampilan proses, melalui fungsi ini dapat diatur penerapan model dari dataset yang dipakai pada proses ini terhadap prediksi label yang akan diterapkan.
5. Tambahkan operator *Performance* yang didapat dari menu **Validation** dan submenu **Performance** untuk mengukur tingkat performa dari model yang diterapkan.
6. Pada operator *Performance* tentukan paramater yang akan digunakan yaitu *Root Mean Square Error* (RMSE).
7. Koneksikan semua perintah tersebut sehingga pada layar tampilan proses terlihat alur tahapan.
8. Setelah dilakukan **Running Process** pada *tools* Rapid Miner, didapatkan hasil prediksi objek testing dengan menggunakan algoritma Regresi Linier yang dapat dilihat pada tabel berikut dibawah ini.

Tabel 4. Hasil Prediksi pada Rapid Miner

NO	Y (observation)	Y (predict) Rapid Miner
1	12239321	12239320
2	12229502	12229501
3	12240724	12240723
4	12232308	12232307
5	12237919	12237918
6	12229502	12229501
7	12229502	12229501
8	12242127	12242126
9	12236516	12236515
10	12233711	12233710

Pengujian performa terhadap model dan algoritma dilakukan dengan maksud mengetahui hasil perhitungan yang dianalisa dan mengukur metode serta algoritma yang

digunakan apakah berfungsi dengan baik atau tidak. Hasil nilai RMSE berdasar pada aplikasi Rapid Miner yang didapat dari pengujian data testing adalah sebagai berikut.

PerformanceVector

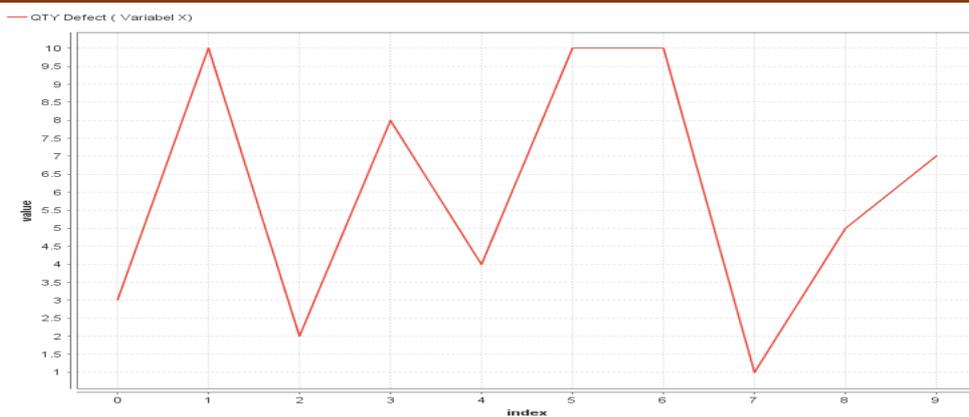
```
PerformanceVector:  
root_mean_squared_error: 0.483 +/- 0.000
```

Persamaan Regresi Linier sederhana tersebut kemudian diimplementasikan dalam memprediksi pada data *testing*. Secara umum penerapan persamaan Regresi Linear tersebut diaplikasikan untuk memprediksi 10 *record* dataset sebagai data uji (*testing*) yang sudah ditentukan. Dari perhitungan yang dilakukan secara manual dan juga dibandingkan dengan proses pada aplikasi Rapid Miner, hasil yang ditunjukkan tidak memiliki perbedaan yang signifikan, dengan kata lain baik dari perhitungan manual maupun yang diproses dalam aplikasi menunjukkan kemiripan hasil. Dibawah ini adalah tabel perbandingan antara hasil perhitungan manual dan yang ada pada aplikasi Rapid Miner terhadap variabel Y.

Tabel 5. Perbandingan Manual dan Rapid Miner

NO	X	Y (observation)	Y (predict) Manual	Y (predict) Rapid Miner
1	3	12239321	12239321	12239320
2	10	12229502	12229502	12229501
3	2	12240724	12240724	12240723
4	8	12232308	12232308	12232307
5	4	12237919	12237919	12237918
6	10	12229502	12229502	12229501
7	10	12229502	12229502	12229501
8	1	12242127	12242127	12242126
9	5	12236516	12236516	12236515
10	7	12233711	12233711	12233710

Sedangkan untuk perbandingan nilai variabel Y aktual dari data *testing* (observasi) dengan nilai variabel Y prediksi pada aplikasi Rapid Miner dapat dilihat melalui grafik berikut.



Gambar 3. Grafik Perbandingan Nilai (Y) Observasi dengan (Y) Prediksi

Secara umum dapat dilihat melalui grafik diatas, dapat dikatakan bahwa nilai prediksi memiliki rentang yang sama dan sebagian besar bernilai lebih tinggi dari nilai Y aktual (observasi). Gap ini nantinya yang akan dievaluasi menggunakan metode RMSE untuk melihat seberapa besar nilai *error* tersebut.

Selanjutnya pada evaluasi model ini adalah dengan mencari nilai *root mean squared error* atau RMSE, melalui aplikasi Rapid Miner didapatkan nilai RMSE sebesar 0,483 dengan standar deviasi +/- 0,0 seperti berikut.

root_mean_squared_error

root_mean_squared_error: 0.483 +/- 0.000

KESIMPULAN DAN REKOMENDASI

Berdasarkan hasil yang penelitian yang dilakukan, kesimpulan yang didapat yaitu sebagai berikut:

1. Pemanfaatan data produk dan *defect* dapat digunakan dalam melakukan proses tahapan data mining dan pemodelan untuk memprediksi jumlah produk *defect* disuatu waktu.
2. Penerapan model persamaan algoritma Regresi Linear sederhana dapat diimplementasikan dimana hasilnya juga menunjukkan sebuah wawasan baru bagi kebutuhan prediksi terhadap jumlah produk *defect*.

3. Model persamaan Regresi Linear sederhana setelah dibandingkan hasil perhitungan secara aktual (observasi) dan juga dengan aplikasi Rapid Miner secara umum menunjukkan kemiripan hasil. Evaluasi dan pengujian performa menggunakan aplikasi Rapid Miner juga dapat menghasilkan gambaran yang relevan dengan skenario yang dimodelkan. Nilai RMSE juga didapat saat melakukan evaluasi performa model yang diterapkan, dengan nilai RMSE sebesar 0,483 dengan standar deviasi $\pm 0,0$.

Beberapa rekomendasi dalam penelitian ini untuk pengembangan lebih lanjut antara lain yaitu:

1. Perbanyak data dalam rentang bulan lebih panjang yang akan digunakan dalam penelitian selanjutnya agar dapat meningkatkan nilai akurasi dari model prediksi terhadap jumlah produk *defect* dimasa mendatang.
2. Pengembangan model dengan menggabungkan beberapa metode dan algoritma yang dapat digunakan dalam memprediksi sesuatu sehingga diharapkan memberikan hasil yang lebih variatif dan bisa dimanfaatkan dalam pengambilan kebijakan dan strategi pengendalian jumlah produk *defect* dan meningkatkan nilai RSME dari model prediksi yang diharapkan.

REFERENSI

- Andini, T. I., Witanti, W., (2016), Prediksi Potensi Pemasaran Produk Baru dengan Metode Naïve Bayes Classifier dan Regresi Linear, pp. 27–32.
- Fengki, F., (2018). Implementasi Regresi Linear Untuk Memprediksi Lama Waktu Pengiriman Catering Kepada Konsumen Studi Kasus Home Catering Malang, *Cent. Libr. Maulana Malik Ibrahim State Islam. Univ. Malang*.
- Fikri, A., (2013). Penerapan Data Mining Untuk Mengetahui Tingkat Kekuatan Beton Yang Dihasilkan Dengan Metode Estimasi Menggunakan Linear Regression, *Fak. Ilmu Komput. UDINUS*, pp. 1–12.
- Gaol, I. L. L., Sinurat, S., and Siagian, E. R., (2019). Implementasi Data Mining Dengan Metode Regresi Linear Berganda Untuk Memprediksi Data Persediaan Buku Pada PT. Yudhistira Ghalia Indonesia Area Sumatera Utara, *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 130–133, doi:

10.30865/komik.v3i1.1579.

- Haryati, S., Sudarsono, A., and Suryana, E., (2018), Implementasi Data Mining untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5, *J. Media Infotama*, vol. 11, no. 2, pp. 130–138.
- Herwanto, H. W., Widiyaningtyas, T., and Indriana, P., (2019). Penerapan Algoritme Linear Regression untuk Prediksi Hasil Panen Tanaman Padi, *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 8, no. 4, p. 364, doi: 10.22146/jnteti.v8i4.537.
- Panggabean, D. S. O., Buulolo, E., and Silalahi, N., (2020). Penerapan Data Mining Untuk Memprediksi Pemesanan Bibit Pohon Dengan Regresi Linear Berganda, *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 1, p. 56, doi: 10.30865/jurikom.v7i1.1947.
- Rahayu, E., Parlina, I., and Siregar, Z. A. (2022). Penerapan Algoritma Regresi Linier Berganda Pada Estimasi Penjualan Sepeda Motor Application of Multiple Linear Regression Algorithm for Motorcycle Sales Estimation, vol. 1, no. 1.
- Ramadhan P. S., and Safitri, N., (2019), Penerapan Data Mining Untuk Mengestimasi Laju Pertumbuhan Penduduk Menggunakan Metode Regresi Linier Berganda Pada BPS Deli Serdang, *Sains dan Komput.*, vol. 18, no. 1, pp. 55–61.
- Rizky, F., Syahra, Y., and Mariami, I., (2019). Implementasi Data Mining Untuk Memprediksi Target Pemakaian Stok Barang Menggunakan Metode Regresi Linier Berganda, *Sains dan Komput.*, vol. 18, no. 2, pp. 167–175.
- Sagita, NR, Octaviani, V, & Indria, I (2023). Pengaruh Kualitas Barang Terhadap Keputusan Pembelian Pada Aplikasi Shopee. *Jurnal Multidisiplin Dehasen ...*, jurnal.unived.ac.id, <https://jurnal.unived.ac.id/index.php/mude/article/view/4291>
- Sugiyono, (2016). Dokumen Karya Ilmiah | Skripsi | Prodi Teknik Informatika - S1 | FIK | UDINUS, *Fik*, vol. 1, no. 1, pp. 1–2.
- Suyanto, *Data Mining*. (2017), Yogyakarta: Informatika.
- Wibisono, A., Rofik, M., and Purwanto, E., (2019). Penerapan Analisis Regresi Linier Berganda dalam Penyelesaian Skripsi Mahasiswa, *J. ABDINUS J. Pengabdian Nasant.*, vol. 3, no. 1, p. 30, doi: 10.29407/ja.v3i1.13512.
- Wulandari, R. T., (2017). *Data Mining*. Yogyakarta: Gava Media.