

Model *Multiple Linear Regression* Untuk Prediksi Durasi Penginstalan *Equipment Mechanical* Dengan Evaluasi MAE dan RMSE Pada PT Nusareka Prima Engineering

Artamian Katarina Manik^{1*}, Hidayanti Murtina², Aryo Tunjung Kusumo³

^{1,2,3} Sistem Informasi, Universitas Bina Sarana Informatika, Indonesia

^{1*}artamnk@gmail.com, ²hidayati.hym@bsi.ac.id, ³aryo.atk@bsi.ac.id

Abstrak: Memperkirakan durasi instalasi *equipment mechanical* merupakan elemen krusial dalam manajemen proyek konstruksi karena berdampak langsung pada penjadwalan, alokasi sumber daya, dan efisiensi biaya. Kesalahan dalam estimasi waktu dapat memicu keterlambatan penyelesaian proyek dan lonjakan biaya operasional. Oleh sebab itu, dibutuhkan pendekatan yang lebih objektif untuk memprediksi durasi berdasarkan data historis. Penelitian ini bertujuan untuk mengimplementasikan metode Regresi Linier Berganda dalam memprediksi waktu penginstalan *equipment mechanical* di PT Nusareka Prima Engineering. Variabel independen yang dianalisis mencakup jumlah pekerja, tingkat kompleksitas pekerjaan, kondisi area kerja, durasi proses *welding*, serta keterlambatan material, dengan durasi penginstalan *equipment mechanical* sebagai variabel dependen. Tahapan pelaksanaan penelitian ini meliputi identifikasi masalah, tinjauan pustaka, pengumpulan data historis, *data preprocessing*, pemisahan dataset menjadi 80% data *training* dan 20% data *testing*, pembuatan model menggunakan perangkat lunak RapidMiner, serta evaluasi performa model melalui *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE). Hasil analisis menunjukkan bahwa model Regresi Linier Berganda mendapatkan nilai RMSE sebesar 1,141 dan MAE sebesar 0,910. Nilai tersebut membuktikan bahwa tingkat kesalahan prediksi relatif rendah dan model mampu menghasilkan estimasi yang mendekati kondisi aktual lapangan. Oleh karena itu, model ini dinilai layak untuk dijadikan alat bantu pengambilan keputusan dalam penyusunan jadwal, perencanaan sumber daya, dan pengendalian proyek, yang pada akhirnya dapat meningkatkan efektivitas proses perencanaan konstruksi di PT Nusareka Prima Engineering.

Kata Kunci: Regresi Linier Berganda; Prediksi Durasi; *equipment mechanical*; MAE; RMSE.

Abstract: Estimating the duration of mechanical equipment installation is a crucial element in construction project management, as it directly impacts scheduling, resource allocation, and cost efficiency. Errors in time estimation can lead to project completion delays and increased operational costs. Therefore, a more objective approach based on historical data is required to predict these durations. This study aims to implement the Multiple Linear Regression method to predict mechanical equipment installation times at PT Nusareka Prima Engineering. The

independent variables include the number of workers, job complexity, work area conditions, welding duration, and material delays, while the mechanical equipment installation duration serves as the dependent variable. The research process involved problem identification, literature review, historical data collection, data preprocessing, splitting the dataset into 80% training data and 20% testing data, model development using RapidMiner software, and model performance evaluation via Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The analysis results indicate that the Multiple Linear Regression model achieved an RMSE of 1.141 and an MAE of 0.910. These values demonstrate a relatively low level of prediction error, indicating that the model is capable of generating estimates that closely reflect actual field conditions. Consequently, the model is considered a viable decision-support tool for scheduling, resource planning, and project control, ultimately enhancing the effectiveness of the construction planning process at PT Nusareka Prima Engineering.

Keywords: Multiple Linear Regression; Duration Prediction; Mechanical Equipment; MAE; RMSE.

1. PENDAHULUAN

Teknologi berkembang disegala bidang dan salah satu yang mengalami perkembangan pesat adalah bidang *machine learning*, yang merupakan salah satu cabang *artificial intelligent* (AI) yang berfokus pada kemampuan sistem komputer untuk belajar dari data. Penggunaan machine learning yang cukup banyak dimanfaatkan adalah dalam melakukan prediksi dan salah satu metodenya adalah regresi linier berganda. Regresi linier dimanfaatkan untuk memprediksi nilai Perusahaan, peningkatan omzet sampai melakukan perencanaan strategi bisnis. Namun belum semua perusahaan memanfaatkan teknologi machine learning ini, salah satunya Adalah PT. Nusareka Prima Engineering.

Di dunia industri konstruksi khususnya dalam melakukan instalasi mesin, durasi pekerjaan sangat penting bagi keberhasilan proyek. Durasi pekerjaan sendiri dipengaruhi oleh beberapa proses teknis seperti pemasangan, pengelasan, dan uji coba yang dipengaruhi oleh berbagai faktor. Di PT. Nusareka Prima Engineering estimasi durasi pekerjaan masih dilakukan secara manual menggunakan pengalaman, yang berpotensi meningkatkan ketidaktepatan jadwal. Kesalahan dalam melakukan prediksi dapat memberikan efek domino terhadap perusahaan. Selain mengakibatkan penundaan jadwal penyelesaian proyek, juga dapat meningkatkan biaya operasional, penurunan nilai equipment, dan menurunnya kepercayaan klien terhadap PT. Nusareka Prima Engineering. Untuk itu penggunaan metode regresi linier dalam memprediksi durasi pekerjaan pada PT. Nusareka Prima Engineering dapat diterapkan.

Pada penelitian sebelumnya Regresi Linier Berganda dianggap sebagai metode yang populer karena kemampuannya untuk mensimulasikan hubungan antar variabel dan menghasilkan hasil yang konsisten dengan kenyataan[1][2]. Namun, banyak penelitian sebelumnya yang berfokus pada proyek konstruksi secara umum atau hanya menggunakan variabel seperti jumlah pekerja dan tingkat produktivitas pekerja. Selain itu, Penelitian yang berfokus pada estimasi durasi pekerjaan dengan menganalisis variabel-variabel seperti jumlah karyawan, kondisi area kerja,, durasi proses pengelasan, keterlambatan material masih sangat terbatas. Kondisi ini menunjukkan adanya banyak aspek penelitian yang perlu dikaji lebih lanjut. Penelitian ini menggunakan metode Regresi Linier Berganda juga dapat menjelaskan hubungan antara satu variabel dependen dan beberapa variabel independen secara simultan. Selain memiliki proses yang sederhana dan mudah diintegrasikan, metode ini juga dapat diterapkan pada data numerik yang memiliki hubungan linier antar variabel sehingga hasil prediksi dapat dijelaskan secara statistik[3]. Terdapat pula penelitian yang membandingkan akurasi pada metode Regresi

Linier Berganda dan metode *Decision Tree* untuk prediksi BMI pada dataset Asthma. Hasil penelitian menunjukkan bahwa Regresi Linier Berganda menghasilkan hasil yang lebih akurat dibandingkan dengan Decision Tree berdasarkan nilai RMSE untuk berbagai set analisis data. Temuan tersebut menunjukkan bahwa Regresi Linier Berganda memberikan potensi prediksi dan memenuhi persyaratan ini secara jauh lebih baik[4][5]. Meskipun penelitian mengenai prediksi durasi proyek telah dilakukan menggunakan berbagai metode, seperti pohon keputusan, metode regresi linier berganda (MLR) tetap menjadi metode yang sangat relevan dalam praktik industri karena interpretasi koefisien variabelnya yang transparan. Oleh karena itu, model ini dapat memberikan prediksi yang akurat mengenai durasi penginstalan Equipm

Perbedaan antara penelitian ini dan penelitian sebelumnya didasarkan pada objek penelitian variabel yang digunakan, dan penggunaan data dari proyek PT Nusareka Prima Engineering. Selain itu, model di evaluasi menggunakan nilai ent mechanical dalam proyek konstruksi.

Perbedaan antara penelitian ini dan penelitian sebelumnya didasarkan pada objek penelitian variabel yang digunakan, dan penggunaan data dari proyek PT Nusareka Prima Engineering. Selain itu, model di evaluasi menggunakan nilai Mean Absolute Error (MAE) dan Mean Squared Error (RMSE) sehingga akurasi model dapat ditentukan secara kuantitatif. Nilai error yang rendah menunjukkan bahwa hasil prediksi sesuai dengan data proyek[6][7][8].

Tujuan dari penelitian yang penulis lakukan adalah menerapkan metode regresi linear berganda untuk memprediksi durasi pengintalan equipment mechanical di PT Nusareka Prima Engineering serta mengevaluasi efisiensi model menggunakan metrik Mean Square Error (RMSE) dan Mean Absolute Error (MAE), guna mengukur tingkat kesalahan prediksi dan memastikan model memiliki tingkat prediktabilitas dan memastikan bahwa model tersebut layak digunakan dalam proyek di PT Nusareka Priima Engineering. Adapun variabel yang penulis gunakan dalam penelitian ini antara lain jumlah tenaga kerja, kondisi area, kompleksitas pekerjaan, keterlambatan material, durasi pemasangan equipment mechanical. Diharapkan dengan adanya penelitian ini dapat membantu Perusahaan dalam melakukan prediksi agar lebih efektif dan akurat sehingga proses perencanaan, *monitoring*, dan pengendalian waktu pekerjaan dapat terbantu dan dapat meminimalkan risiko keterlambatan proyek konstruksi.

2. METODE PENELITIAN

Berikut ini tahapan penelitian yang penulis gunakan untuk memprediksi durasi pemasangan equipment mechanical pada PT Nusareka Prima Engineering.



Gambar 1. Tahapan Penelitian

Berikut ini penjelasan dari gambar 1 tahapan penelitian yang penulis gunakan untuk memprediksi durasi pemasangan equipment mechanical pada PT Nusareka Prima Engineering, antara lain:

A. Identifikasi Masalah

Pada tahap ini, penulis mengidentifikasi permasalahan yang terjadi selama proses instalasi *equipment mechanical* di PT Nusareka Prima Engineering. Berdasarkan hasil observasi, penulis menyimpulkan bahwa estimasi durasi pekerjaan masih dilakukan berdasarkan pengalaman dan perkiraan, yang dapat mengakibatkan ketidaksesuaian antara waktu aktual yang dihabiskan untuk pekerjaan dan waktu yang direncanakan. Kondisi tersebut dapat menyebabkan keterlambatan proyek, ketidak efisiensi dalam penggunaan sumber daya sehari-hari, dan peningkatan biaya operasional. Karena itu, penulis mengusulkan penggunaan metode Regresi Linier Berganda untuk membantu memprediksi durasi penginstalan *equipment mechanical* secara lebih akurat dengan menggunakan data riwayat proyek.

B. Studi Literatur

Pada tahap studi literatur, penulis mempelajari berbagai referensi yang berkaitan dengan penelitian, seperti jurnal ilmiah, buku, artikel penelitian dan penelitian terdahulu yang membahas Regresi Linier Berganda, *Supervised Learning*, RapidMiner, serta metode evaluasi prediksi. Selain itu penulis juga mempelajari data historis proyek yang dimiliki Perusahaan untuk memahami karakteristik data yang akan digunakan dalam penelitian.

C. Pengumpulan Data

Tahap pengumpulan data dilakukan dengan mengumpulkan data historis dari proyek penginstalan *Equipment mechanical* yang diperoleh dari PT Nusareka Prima Engineering. Pengumpulan data ini tidak terbatas hanya satu jenis *Equipment* saja melainkan juga mencakup berbagai jenis *equipment mechanical* baik rotatik maupun statik seperti pump, vacuum blower, agitator, dedusting station. Semua jenis equipment tersebut dianggap sebagai objek penelitian karena proses instalasinya melibatkan kegiatan pemasangan sebelum tahap komisioning. Variabel data yang digunakan meliputi jumlah pekerja, kompleksitas pekerjaan, kondisi area kerja, durasi proses welding, keterlambatan material, durasi penginstalan *equipment*. Penulis menggunakan data historis tersebut sebagai sumber utama untuk membangun model prediksi karena data tersebut menunjukkan kondisi proyek yang telah diselesaikan sebelumnya.

D. Data Preprocessing

Setelah Data berhasil dikumpulkan, penulis melakukan tahap *data preprocessing* untuk meningkatkan kualitas data sebelum digunakan dalam proses pemodelan. Tahapan ini meliputi pembersihan data (*data cleaning*) untuk menghapus data yang tidak lengkap atau tidak konsisten, seleksi data untuk memilih variabel yang relevan, transformasi data ke dalam format yang terstruktur serta pembagian data menjadi data *training* dan *testing* untuk proses pembangunan dan pengujian model prediksi. Tahapan ini bertujuan agar data yang digunakan memiliki kualitas yang baik sehingga dapat menghasilkan model prediksi yang lebih akurat[9]. Penulis menggunakan pembagian data 80:20, pembagian ini dipilih karena memberikan proporsi data yang lebih besar untuk proses pelatihan model sehingga hubungan antar variabel dapat dipelajari dengan baik, sementara data *testing* digunakan untuk mengevaluasi kemampuan prediksi model terhadap data yang belum pernah digunakan sebelumnya. Ratio 80:20 merupakan salah satu proporsi yang umum digunakan dalam pemodelan prediktif dan *Machine Learning*[15].

E. Metode Regresi Linier Berganda

Pada tahap ini, penulis menerapkan metode Regresi Linier Berganda untuk membangun model prediksi durasi penginstalan *equipment mechanical*. Metode ini digunakan untuk menganalisis hubungan antara dua atau lebih variabel bebas (independen) terhadap satu variabel terikat (dependen)[8][10][11]. Metode yang digunakan sebagai variabel independen yaitu jumlah tenaga kerja, kompleksitas pekerjaan, kondisi area kerja, durasi proses *welding*, keterlambatan material sedangkan untuk variabel dependen berupa durasi penginstalan *equipment mechanical*. Melalui metode ini, penulis dapat mengetahui pengaruh masing-masing variabel terhadap durasi pekerjaan sekaligus menghasilkan model yang dapat digunakan untuk melakukan prediksi pada data baru. Persamaan dari Regresi Linier Berganda adalah sebagai berikut[12].

$$Y = a + \beta_1 x_1 + \beta_2 x_2 + \dots \beta_n x_n \quad (1)$$

Keterangan,

Y = variabel tak bebas (nilai variabel yang akan diprediksi)

a = konstanta

β = nilai koefisien regresi

x = variabel bebas

F. Evaluasi Prediksi

Tahap evaluasi dilakukan untuk mengukur tingkat akurasi model yang telah dibangun. Penulis menggunakan metode *Root Mean Squared Error* (RMSE) dan *Mean Absolute Error* (MAE) sebagai matrix evaluasi. MAE digunakan untuk menghitung rata-rata kesalahan absolut antara nilai saat ini dan nilai prediksi, sedangkan RMSE digunakan untuk menghitung kesalahan absolut dengan memberikan penalti yang lebih besar untuk kesalahan yang tinggi. Hasil evaluasi digunakan untuk memahami beberapa model Regresi Linier Berganda yang baik dalam memprediksi durasi penginstalan *equipment mechanical*.

1) Mean Absolute Error (MAE)

MAE adalah metode evaluasi yang digunakan untuk menghitung perbedaan absolut antara nilai saat ini dan nilai prediksi dalam model regresi Pada garis linier berganda, semakin kecil nilai MAE yang diperoleh, maka semakin baik kemampuan model dalam melakukan prediksi karena hasil prediksi semakin mendekati data sebenarnya.

Rumus *Mean Absolute Error* (MAE)[13]:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (2).$$

Keterangan:

MAE = *Mean Absolute Error*

n = jumlah data

Y_i = nilai aktual

$\hat{Y}_i$ = nilai prediksi

$|Y_i - \hat{Y}_i|$ = nilai absolut *error* antara data aktual dan prediksi

2) Root Mean Square Error (RMAE)

RMSE adalah ukuran kesalahan prediksi rata-rata dalam skala asli variabel dependen. Dalam regresi linier berganda, RMSE dihitung untuk mengetahui seberapa jauh prediksi model ($\hat{Y}$) berbeda dari nilai aktual (Y). Nilai RMSE yang lebih kecil menunjukkan model yang lebih akurat.

Rumus *Root Mean Square Error* (RMAE)[14]:

$$RMSE = \sqrt{\sum_{i=1}^n \cdot (y_i - \hat{y}_i)^2} \quad (3)$$

Keterangan:

Keterangan

y_i = Nilai aktual data ke- i

$\hat{y}_i$ = Nilai prediksi data ke- i

n = Jumlah total observasi

G. Hasil Prediksi

Pada tahapan ini, penulis menganalisis hasil prediksi yang diperoleh menggunakan model Regresi Linier Berganda. Hasil prediksi tersebut dibandingkan dengan data aktual untuk menentukan tingkat kesesuaian model dengan kondisi sebenarnya. Berdasarkan hasil tersebut, penulis dapat memahami kemampuan model dalam memperkirakan durasi penginstalan *equipment mechanical* berdasarkan variabel-variabel yang digunakan dalam penelitian.

H. Kesimpulan

Berdasarkan hasil penelitian, metode Regresi Linier Berganda dapat digunakan untuk memprediksi durasi penginstalan *equipment mechanical* di PT Nusareka Prima Engineering menggunakan data riwayat proyek. Metode penelitian dilakukan melalui identifikasi masalah, studi literatur, pengumpulan data, data preprocessing, metode regresi linier berganda, dan evaluasi menggunakan MAE dan RMSE. Temuan penelitian menunjukkan bahwa faktor-faktor seperti jumlah pekerja, kompleksitas pekerjaan, kondisi area kerja, lamanya proses pengelasan, dan keterlambatan material terhadap lamanya pengerjaan. Model yang dihasilkan mampu memberikan estimasi durasi penginstalan *equipment mechanical* yang mendekati kondisi aktual, yang dapat membantu perusahaan dalam perencanaan jadwal, pengelolaan sumber daya, dan pengambilan keputusan proyek secara lebih efektif.

3. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Penulis melakukan pengumpulan data penelitian yang merupakan data primer Perusahaan dalam melakukan pengerjaan pemasangan *equipment mechanical* di tahun 2025. Dari hasil pengumpulan data diperoleh sebanyak 80 *record* seperti yang terlihat pada tabel 2 yang nantinya akan penulis gunakan sebagai bahan penelitian.

Tabel 2. Dataset

Nomor	Jumlah tenaga kerja	Kompleksitas kerja	Kondisi area kerja	Durasi proses welding	Keterlambatan material	Durasi penginstalan equipment
1	3	2	3	6	2	10
2	6	4	5	7	3	15
3	6	3	5	6	1	10
4	5	2	4	6	4	13
5	6	2	4	5	2	9
...	...	...	...	...	...	...
...	...	...	...	...	...	...
76	5	2	4	2	1	12
77	7	4	4	7	3	14
78	4	2	2	4	5	12

79	4	4	3	5	4	13
80	6	2	2	3	7	11

Berikut ini merupakan penjelasan pada tabel 2, yaitu:

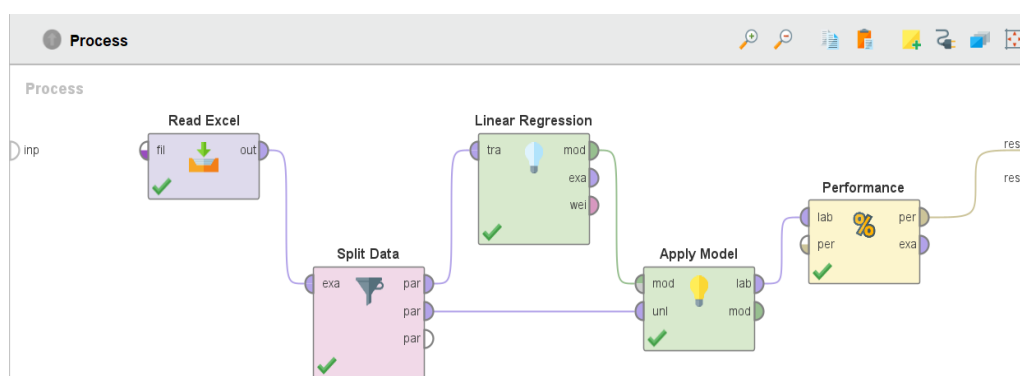
- 1) Jumlah tenaga kerja = Jumlah tenaga yang terlibat dalam proses pemasangan (orang)
- 2) Kompleksitas pekerjaan = Skala 1-5, semakin tinggi nilai semakin kompleks pekerjaan
- 3) Kondisi area kerja = Skala 1-5, semakin tinggi nilai menunjukkan area kerja semakin sulit.
- 4) Durasi proses *welding* = Lama proses pengelasan (Hari)
- 5) Keterlambatan material = lama keterlambatan material (hari)
- 6) Durasi penginstalan *equipment* = Lama waktu pemasangan hingga pekerjaan selesai

B. Data Preprocessing

Penulis proses preprocessing terhadap 80 data yang telah penulis kumpulkan dengan melakukan data *cleaning* dan penyeleksian variabel yang akan digunakan. Hasilnya penulis mendapatkan 80 data yang sudah dibersihkan dan menetapkan variabel Jumlah tenaga kerja, Kompleksitas pekerjaan, Kondisi area kerja, Durasi proses *welding* dan Keterlambatan material sebagai variabel (X), sedangkan Durasi penginstalan *equipment* sebagai variabel (Y).

C. Metode Regresi Linier Berganda

Pada penerapan metode regresi linier berganda dan pengujian akurasi nilai MAE dan RMSE penulis menggunakan tools bantuan RapidMiner dengan hasil seperti yang terlihat pada gambar 2.



Gambar 2. Penerapan RapidMiner Metode Regresi Linier

Berikut ini penjelasan untuk penerapan RapidMiner pada gambar 2, antara lain:

1) Dataset Input Testing dan Training

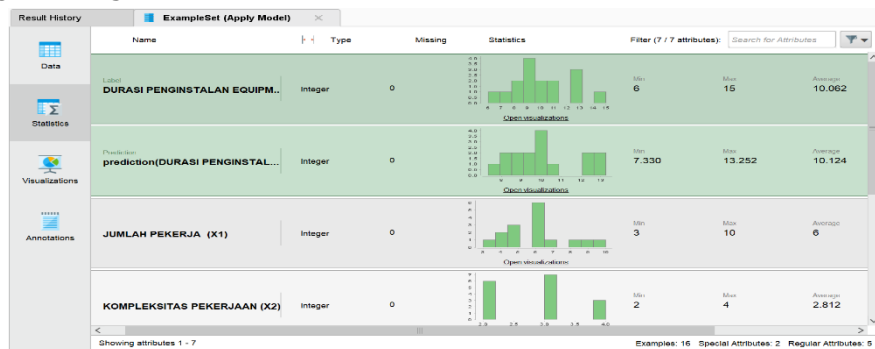
Dataset dibagi menjadi dua, yaitu *training* dan *testing*. Data pada penelitian ini dibagi menjadi 80% sebagai data *training* dan 20% sebagai data *testing*. Jumlah data yang diinput 80 data, setelah dilakukan pengujian *training* dan *testing* menghasilkan 64 data.

2) Regresi Linier

Penulis menggunakan objek linear regression untuk menerapkan metode regresi linier berganda yang dihubungkan dengan objek split data yang sebelumnya sudah diatur pembagian datanya. Pada tools RapidMiner penggunaan objek Linear Regression sudah terdapat rumus 1.

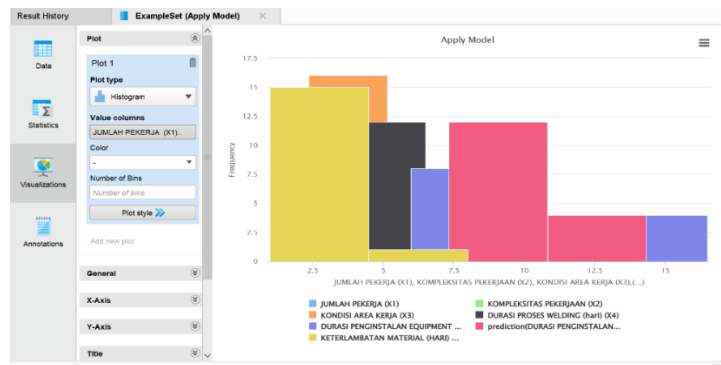
3) Apply Model

Apply model merupakan tahap terakhir Dimana model data mining regresi linier diterapkan berdasarkan proses-proses sebelumnya, termasuk penggunaan data *testing*, data *training* dan regresi linier.



Gambar 3. Tampilan Statistics ExampleSet

Pada gambar 3. Tabel statistik deskriptif menunjukkan bahwa dataset terdiri dari 16 examples dengan 7 atribut, yang terdiri atas 2 atribut khusus dan 5 atribut reguler. tidak ditemukan adanya *missing values* dalam dataset, sehingga seluruh data dapat digunakan langsung dalam proses pemodelan. Nilai rata-rata untuk durasi penginstalan *equipment* aktual adalah 10.062, yang sangat mendekati nilai rata-rata hasil prediksi model sebesar 10.124. Hal ini memberikan indikasi awal bahwa model memiliki bias yang rendah dalam memprediksi nilai rata-rata populasi.



Gambar 4. Tampilan Histogram

Gambar 4 mengilustrasikan visualisasi *Apply Model* yang digunakan untuk menentukan distribusi data secara sederhana. Histogram ini menggambarkan distribusi frekuensi variabel selama proses tersebut. Melalui fitur visualisasi ini, verifikasi awal terhadap data yang dihasilkan oleh model prediksi dilakukan sebelum analisis statistik yang lebih mendalam yang berfokus pada penggunaan metrik MAE dan RSME untuk memastikan akurasi kinerja model secara kuantitatif.

4) Performance

Penulis menggunakan Objek *Performance* untuk melihat hasil akurasi dari prediksi durasi penginstalan *equipment* menggunakan algoritma regresi linier dengan memberikan gambaran tentang seberapa baik model tersebut dapat memprediksi durasi penginstalan. Evaluasi performa model regresi linier dilakukan menggunakan dua metrik, yaitu *Mean*

Absolute Error (MAE) dan *Root Mean Square Error* (RMSE). MAE mengukur rata-rata selisih absolut antara nilai aktual dan prediksi.

MAE didefinisikan sebagai:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (2).$$

RMSE didefinisikan sebagai:

$$RMSE = \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

Di mana y_i adalah nilai aktual, $\hat{y}_i$ adalah nilai prediksi, n adalah jumlah data. Perhitungan MAE dan RMSE dilakukan menggunakan operator *performance* (Regression) yang tersedia di RapidMiner. Adapun langkah-langkahnya sebagai berikut:

1. Model regresi linier diterapkan pada data uji (*testing data*).
2. Output prediksi dari model dihubungkan ke port input operator *performance* (Regression).
3. Operator *performance* (Regression) secara otomatis menghitung nilai MAE dan RMSE berdasarkan persamaan (2) dan (3).
4. Hasil perhitungan MAE dan RMSE yang dihasilkan kemudian dianalisis untuk mengevaluasi akurasi model

Berikut hasil perhitungan menggunakan MAE dan RMSE terlihat pada gambar dibawah ini.

PerformanceVector

```
PerformanceVector:  
root_mean_squared_error: 1.141 +/- 0.000  
absolute_error: 0.910 +/- 0.689
```

Gambar 5. Hasil *Performance vector*

Dari hasil nilai *performance vector*, hasil penilaian *Root Mean Square Error* (RMSE) dan *Absolute Error* yang dilakukan dengan memanfaatkan operator *performance* pada RapidMiner, didapatkan nilai sebesar 1.141+/- 0.0 untuk RMSE, 0.910 +/- 0.0689 untuk *Absolute error*.

4. KESIMPULAN

Penelitian menggunakan populasi data primer penginstalan *equipment mechanical* pada PT. Nusareka Prima Engineering selama tahun 2025 sebanyak 80 data. Dengan penggunaan variabel Jumlah tenaga kerja, Kompleksitas pekerjaan, Kondisi area kerja, Durasi proses *welding* dan Keterlambatan material sebagai variabel (X), sedangkan Durasi penginstalan *equipment* sebagai variabel (Y). Hasil dari penelitian ini Adalah dapat digunakannya metode regresi linier berganda dalam memprediksi durasi penginstalan *equipment mechanical* pada PT. Nusareka Prima Engineering dengan menggunakan 80% data *training* dan 20% data *testing*. Serta hasil evaluasi menggunakan RMSE dan MAE menunjukkan nilai RMSE sebesar 1,141 dan MAE sebesar 0,910, yang menunjukkan bahwa model memiliki tingkat kesalahan prediksi rendah sehingga layak digunakan untuk membantu memperkirakan durasi penginstalan *equipment mechanical*.

5. UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada PT Nusareka Prima Engineering yang telah memberikan izin, kesempatan, serta dukungan selama proses pengumpulan data dan pelaksanaan penelitian.

6. REFERENCES

- [1] A. Rasuli, R. Kusuma Serli, K. Indah Panca Ningtyas, and S. Yozha Putri Kinanti, "Prediksi Stok Produk Susu Pada Pt Greenfields Dairy Indonesia Menggunakan Metode Regresi Linier," *CONTENT Comput. Netw. Technol.*, vol. 5, no. 1, pp. 23–28, 2025, [Online]. Available: <http://jurnal.bsi.ac.id/index.php/conten>
- [2] Siswo Adiguno, Yohanni Syahra, and Milfa Yetri, "5331-8141-1-Pb," *J. Sist. Inf. Tgd*, vol. 1, pp. 275–281, 2022.
- [3] M. O. Sanni-anibire, "Developing a machine learning model to predict the construction duration of tall building projects," vol. 4, no. 1, pp. 22–36, 2021.
- [4] A. T. Nurani, A. Setiawan, B. Susanto, D. Salatiga, and J. Tengah, "Perbandingan Kinerja Regresi Decision Tree dan Regresi Linear Berganda untuk Prediksi BMI pada Dataset Asthma," vol. 6, no. 1, pp. 34–43, 2023.
- [5] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, pp. 1–24, 2021, doi: 10.7717/PEERJ-CS.623.
- [6] Y. Tambunan, Y. Yohannes, and S. Sipayung, "Pemodelan Tren Pendapatan Game Gacha Menggunakan Regresi Linier," vol. 6, no. 1, pp. 49–58, 2026.
- [7] D. Arifuddin, K. Kusrini, and K. Kusnawi, "Perbandingan Performansi Algoritma Multiple Linear Regression dan Multi Layer Perceptron Neural Network dalam Memprediksi Penjualan Obat," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 2, pp. 722–737, 2025, doi: 10.57152/malcom.v5i2.1952.
- [8] S. A. Nur Cahyo and M. T. Sulistyono, "Comparison of Multiple Linear Regression and Random Forest Methods for Predicting National Rice Production in Indonesia," *J. Appl. Informatics Comput.*, vol. 9, no. 6, pp. 3479–3489, 2025, doi: 10.30871/jaic.v9i6.11398.
- [9] J. S. Sebayang and B. Yuniarto, "Perbandingan Model Estimasi Artificial Neural Network Optimasi Genetic Algorithm dan Regresi Linier Berganda," *Media Stat.*, vol. 10, no. 1, p. 13, 2017, doi: 10.14710/medstat.10.1.13-23.
- [10] S. Ningsih and H. H. Dukalang, "Penerapan Metode Suksesif Interval pada Analisis Regresi Linier Berganda," *Jambura J. Math.*, vol. 1, no. 1, pp. 43–53, 2019, doi: 10.34312/jjom.v1i1.1742.
- [11] S. D. Cahyo, A. A. Murtopo, and B. A. Santoso, "Penerapan Metode Regresi Linier untuk Prediksi Jumlah Penumpang Kereta Api," *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 3, pp. 986–993, 2025, doi: 10.31004/riggs.v4i3.2061.
- [12] R. P. Simbolon and S. P. Sipayung, "Penerapan metode regresi linear berganda dalam mengestimasi jumlah penduduk pada Kabupaten Samosir," *Semin. Nas. Inov. Sains Teknol. Inf. Komput.*, vol. 1, no. 2, pp. 426–436, 2024, [Online]. Available: <https://ejournal.ust.ac.id/index.php/SNISTIK/article/view/3660>
- [13] W. Inawati, R. Kurniawan, and Y. A. Wijaya, "Penerapan Aplikasi Rapidminer pada Prediksi Data Penjualan Album K-Pop Menggunakan Metode Regresi Linear," *J. Ris. dan Apl. Mhs. Inform.*, vol. 5, no. 3, pp. 585–592, 2024, doi: 10.30998/jrami.v5i3.10648.
- [14] A. Hurifiani, A. Irma Purnamasari, and I. Ali, "Penerapan Algoritma Regresi Linear Untuk Prediksi Penjualan Alat Tulis Kantor (Atk) Di Bumdes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 266–273, 2024, doi: 10.36040/jati.v8i1.8305.
- [15] V. R. Joseph, "Optimal ratio for data splitting," *Stat. Anal. Data Min.*, vol. 15, no. 4, pp. 531–538, 2022, doi: 10.1002/sam.11583.