

# Aplikasi Machine Learning Untuk Forecasting Nilai Overall Equipment Effectiveness Pada Industri Manufaktur

Aufatus Syakira Mardhatillah<sup>1</sup>, Dina Agustina<sup>2</sup>

<sup>1,2</sup>Prodi Matematika, Fakultas Matematika Ilmu Pengetahuan dan Alam Universitas Negeri Padang (UNP)

## Article Info

### Article history:

Received January 31, 2023

Revised May 11, 2023

Accepted June 30, 2023

### Keywords:

Manufacturing Industry  
Overall Equipment  
Effectiveness (OEE)  
Machine Learning  
Linear Regression  
Mean Absolute Percentage  
Error

### Kata Kunci:

Industri Manufaktur  
Overall Equipment  
Effectiveness (OEE)  
Machine Learning  
Regresi Linear  
Mean Absolute Percentage  
Error

## ABSTRACT

Manufacturing productivity is determined by machine performance. The problem that occurs in production machines is the occurrence of downtime so that the machine does not work optimally. One of the engine performance indicators is OEE. forecasting of the OEE value is needed so that manufactures can take action so they can maintain engine performance. The purpose of this research was to see the results of forecasting OEE value with machine learning linear regression algorithms and see the value of model evaluation with MAPE. The results of the study using training data and testing data obtained a multiple regression model with the variable y is OEE, the intercept coefficient is  $-53,72$ , and slope coefficient of the availability and performance variables is  $0,87$  and  $0,63$ . The model evaluation results are in the range  $<10\%$ , meaning that the accuracy results had excellent forecasting model.

## ABSTRAK

Produktivitas manufaktur ditentukan oleh kinerja mesin. Permasalahan yang terjadi pada mesin produksi adalah terjadinya *downtime* sehingga mesin tidak bekerja secara optimal. Salah satu indikator kinerja mesin yaitu OEE. *Forecasting* terhadap nilai OEE diperlukan agar manufaktur bisa melakukan tindakan sehingga bisa menjaga kinerja mesin. Tujuan penelitian ini untuk melihat hasil *forecasting* nilai OEE dengan *machine learning* algoritma regresi linear dan melihat nilai evaluasi model dengan MAPE. Hasil penelitian menggunakan data *training* dan data *testing* dengan rasio 7:3 diperoleh model regresi berganda dengan variabel y adalah OEE, koefisien *intercept* yaitu  $-53,72$ , dan koefisien *slope* variabel *availability* dan *performance* yaitu  $0,87$  dan  $0,63$ . Hasil evaluasi model berada pada *range*  $<10\%$ , artinya hasil akurasi memiliki model peramalan sangat baik.

This is an open access article under the [CC BY-SA](#) license.



## Penulis pertama/sesuai:

(Aufatus Syakira Mardhatillah)

Prodi Matematika, Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,

Universitas Negeri Padang, Jl. Prof. Dr. Hamka, Air Tawar barat, Padang Utara, Padang, 25171 Padang, Sumatera Barat

Email: [syakiraaufatus@gmail.com](mailto:syakiraaufatus@gmail.com)

## 1. PENDAHULUAN

Industri manufaktur mempunyai peranan penting dalam perekonomian masyarakat di Indonesia. Kartasapoetra mengatakan bahwa industri manufaktur merupakan perusahaan yang memproduksi bahan baku menjadi barang jadi yang memiliki nilai guna lebih tinggi dengan menggunakan mesin produksi dalam skala produksi yang sangat besar [1]. Sehingga mampu menghasilkan produk yang dibutuhkan oleh pasar kemudian dapat diperdagangkan. Produktivitas



manufaktur ditentukan oleh kinerja mesin. Menurut Nakajima [2], peningkatan hasil produksi pada industri manufaktur dapat dilihat dari segi mesin.

Efektivitas mesin mempengaruhi hasil produksi manufaktur. Dalam proses produksi, permasalahan yang terjadi adalah terjadinya *downtime* sehingga mesin tidak bekerja secara optimal yang mengakibatkan proses produksi terganggu. Jika mesin tidak bekerja secara optimal maka target dan hasil tidak sesuai dengan yang diinginkan oleh perusahaan. Agar mesin dapat bekerja secara optimal, perlu dilihat keefektifan kinerja dari suatu mesin berdasarkan *Key Performance Indicator* (KPI).

KPI merupakan indikator kinerja yang bersifat terukur dan dianggap sebagai kunci keberhasilan untuk mencapai tujuan bisnis sebuah perusahaan [3]. Salah satu indikator keberhasilan rantai produksi manufaktur berdasarkan KPI, yaitu *Overall Equipment Effectiveness* (OEE) [4]. Nakajima (1984) mengatakan bahwa OEE merupakan suatu perhitungan yang dilakukan guna menentukan tingkat produktivitas mesin [5]. Salah satu tindakan yang bisa dilakukan perusahaan untuk mengamati kinerja mesin adalah dengan melakukan *forecasting* terhadap nilai OEE. Dengan adanya revolusi 4.0 dan pemanfaatan big data, perusahaan dapat memanfaatkan aplikasi *machine learning* untuk peningkatan produktivitas mesin.

*Machine learning* merupakan suatu aplikasi komputer dan algoritma matematika yang mampu belajar dari data-data yang ada untuk pengambilan keputusan serta dapat melakukan prediksi di masa mendatang [6]. Secara umum algoritma pembelajaran dalam *machine learning* dikategorikan menjadi *supervised learning* merupakan proses pembelajaran di bawah pengawasan *supervisor* dimana algoritma ini belajar dari data berlabel yang biasanya dijadikan *output* model dan *unsupervised learning* merupakan proses pembelajaran dilakukan tanpa pengawasan dimana algoritma ini belajar dari data yang tidak berlabel [7]. Kelebihan dari *machine learning*, yaitu algoritma atau program komputer memiliki kemampuan belajar secara otomatis berdasarkan data yang ada untuk mengetahui karakteristik dari data. Sehingga pemanfaatan *machine learning* dapat digunakan peneliti untuk *forecasting* nilai OEE pada industri manufaktur dengan algoritma regresi linear. Dataset penelitian diperoleh dari manufaktur makanan di Indonesia.

## 2. METODE

Penelitian ini adalah penelitian terapan dan jenis data yang digunakan adalah data sekunder manufaktur makanan di Indonesia selama periode Desember 2020 sampai periode Desember 2021. Untuk menentukan hasil *forecasting* nilai *Overall Equipment Effectiveness* (OEE) pada industri manufaktur dengan pemanfaatan *machine learning* terdapat langkah-langkah dalam menyelesaikan permasalahan penelitian. Maka teknik analisis pada penelitian ini adalah:

1. Melakukan *pre-processing* data.
2. Melakukan penghapusan nilai OEE <0% dan >100% agar dataset bisa digunakan sesuai format yang ada.
3. Melakukan *exploratory data analysis* (EDA) dengan visualisasi data.
4. Melakukan penghapusan *outlier* agar dataset dapat digunakan dalam proses analisis data.
5. Melakukan pemilihan variabel dengan korelasi.
6. Melakukan penerapan regresi linear.
7. Melakukan evaluasi model dengan *Mean Absolute Percentage Error* (MAPE).

### 2.1 Overall Equipment Effectiveness (OEE)

OEE merupakan suatu alat ukur efektivitas mesin dengan *availability*, *performance*, dan *quality* yang merupakan 3 aspek perhitungan OEE [8].

$$OEE = Availability (\%) \times Performance (\%) \times Quality (\%) \quad (1)$$

Keterangan :

*Availability* = Ketersediaan waktu untuk mesin beroperasi

*Performance* = Kinerja mesin dalam menghasilkan produk

*Quality* = Kualitas produk sesuai dengan standar

## 2.2 Pre-processing Data

*Pre-processing* data adalah tahap untuk proses filterisasi data agar dataset yang digunakan menghasilkan nilai *output* yang lebih baik dimana dilakukan *cleaning* terhadap *missing value* dan data duplikat [9]. *Pre-processing* dilakukan agar dataset dapat digunakan dalam proses penelitian.

## 2.3 Exploratory Data Analysis (EDA)

EDA adalah proses menganalisis dan menampilkan data yang bertujuan untuk memperoleh suatu informasi dan mendeteksi outlier pada dataset yang dapat ditampilkan dalam bentuk grafis seperti boxplot dan histogram [10]. Boxplot adalah jenis visualisasi data untuk melihat data *outlier* pada dataset, sedangkan histogram adalah jenis visualisasi data untuk menampilkan tabulasi frekuensi yang digambarkan dengan grafis batangan dari dataset yang ada [11]. EDA dilakukan agar dapat mengetahui karakteristik dari dataset dan melihat *outlier* pada dataset.

## 2.4 Penghapusan Outlier

*Outlier* adalah data yang menyimpang secara ekstrim dari observasi-observasi lainnya pada setiap variabel [12]. Penghapusan *outlier* dilakukan agar dataset dapat digunakan dalam proses analisis data.

## 2.5 Nilai Korelasi

Nilai korelasi bertujuan untuk melihat keterkaitan erat antara variabel bebas dengan variabel takbebas [13]. Nilai korelasi terletak antara  $-1 \leq r \leq 1$ , dimana :

1. Nilai  $r = -1$  menyatakan korelasi negatif (berlawanan arah) yang kuat antara dua variabel.
2. Nilai  $r = 1$  menyatakan korelasi positif (searah) yang kuat antara dua variabel.
3. Nilai  $r = 0$  menyatakan tidak ada korelasi antara dua variabel.

## 2.6 Regresi

Regresi berganda merupakan model untuk melakukan prediksi yang melibatkan lebih dari satu variabel bebas dimana model ini digunakan untuk melihat keterkaitan erat antara satu variabel takbebas (*output*) dengan dua atau lebih variabel variabel bebas (*input*) [14].

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + \varepsilon \quad (2)$$

Keterangan :

$\alpha$  = Koefisien *intercept*

$\beta_i$  = Koefisien *slope* dimana  $i = 1, 2, 3, \dots, n$

$y$  = Variabel takbebas

$x_i$  = Variabel variabel bebas dimana  $i = 1, 2, 3, \dots, n$

## 2.7 Evaluasi Model

Melakukan evaluasi model dengan MAPE menggunakan Program Python. *Mean Absolute Percentage Error* (MAPE) digunakan untuk mengevaluasi kinerja prediksi yang digunakan dengan mengukur tingkat akurasi dari hasil prediksi [15].

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

dengan

$y_i$  = Data aktual ke 1, 2, 3, ..., n

$\hat{y}$  = Nilai prediksi dari variabel y ke 1, 2, 3, .. n

$n$  = Jumlah data/banyaknya observasi



Adapun skala penilaian akurasi MAPE disajikan pada Tabel 1.

<b>Range MAPE</b>	<b>Akurasi Peramalan</b>
< 10%	Peramalan sangat baik
10 – 20%	Peralaman baik
20 – 50%	Peramalan layak
50%	Peramalan tidak akurat

*Sumber* : Lawrance et al (2009)

### 3. HASIL DAN PAMBAHASAN

Penelitian ini menggunakan Program Python yang akan memudahkan dalam melakukan penelitian prediksi nilai OEE pada industri manufaktur.

#### 3.1 Pre-processing Data

*Pre-processing* data dilakukan agar data sesuai dengan format data. Pada tahap ini dilakukan *cleaning* data, seperti *missing value* dan data duplikat.

##### 1. Missing value

Hasil *cleaning* data sebanyak 684 dataset terhadap *missing value* disajikan pada Tabel 2.

No	Observasi Ke-	Availability	Performance	Quality	OEE
1	0	90,29	111,38	99,99	100,55
2	1	89,09	20,78	100,00	18,51
3	2	82,74	106,98	100,00	88,52
4	3	82,57	108,02	99,94	89,14
5	4	77,91	102,08	100,00	79,53
6	5	89,34	112,27	99,99	100,29
7	6	21,11	75,48	99,99	15,93
8	7	-92,43	46,23	100,00	-42,73
9	8	47,98	19,40	100,00	9,31
...	...	...	...	...	...
629	675	51,62	55,48	99,96	28,63
630	676	71,73	51,75	100,00	37,12
631	677	94,01	108,92	99,98	102,37
632	678	82,01	97,59	100,00	80,04
633	679	95,30	4,82	100,00	4,59
634	680	90,87	108,10	100,00	98,23
635	681	91,25	106,75	99,98	97,39
636	682	97,34	111,85	100,00	108,88
637	683	79,25	91,05	100,00	72,15

Berdasarkan Tabel 2, sebanyak 684 dataset yang ada dilakukan proses *cleaning* data terhadap *missing value* sehingga diperoleh 637 dataset yang tidak memiliki *missing value*.

##### 2. Data duplikat

Hasil *cleaning* data sebanyak 637 dataset terhadap data duplikat disajikan pada Tabel 3.

#### **Tabel 3. Dataset Tanpa Data Duplikat**

No	Observasi Ke-	Availability	Performance	Quality	OEE
1	0	90,29	111,38	99,99	100,55
2	1	89,09	20,78	100,00	18,51
3	2	82,74	106,98	100,00	88,52
4	3	82,57	108,02	99,94	89,14
5	4	77,91	102,08	100,00	79,53
6	5	89,34	112,27	99,99	100,29
7	6	21,11	75,48	99,99	15,93
8	7	-92,43	46,23	100,00	-42,73
9	8	47,98	19,40	100,00	9,31
...	...	...	...	...	...
609	675	51,62	55,48	99,96	28,63
610	676	71,73	51,75	100,00	37,12
611	677	94,01	108,92	99,98	102,37
612	678	82,01	97,59	100,00	80,04
613	679	95,30	4,82	100,00	4,59
614	680	90,87	108,10	100,00	98,23
615	681	91,25	106,75	99,98	97,39
616	682	97,34	111,85	100,00	108,88
617	683	79,25	91,05	100,00	72,15

Berdasarkan Tabel 3, sebanyak 637 dataset yang ada dilakukan proses *cleaning* data terhadap data duplikat sehingga diperoleh 617 dataset yang tidak memiliki data duplikat.

### 3.2 Nilai OEE

Hasil penghapusan Nilai OEE<0% dan >100% yang ditampilkan pada Tabel 4.

**Tabel 4. Dataset dengan Nilai OEE**

No	Observasi Ke-	Availability	Performance	Quality	OEE
1	1	89,09	20,78	100,00	18,51
2	6	21,11	75,48	99,99	15,93
3	8	47,98	19,40	100,00	9,31
4	9	62,69	76,25	100,00	47,80
5	12	53,75	98,00	100,00	52,68
...	...	...	...	...	...
214	675	51,62	55,48	99,96	28,63
215	676	71,73	51,75	100,00	37,12
216	678	82,01	97,59	100,00	80,04
217	679	95,30	4,82	100,00	4,59
218	683	79,25	91,05	100,00	72,15

Berdasarkan Tabel 4, sebanyak 617 dataset yang ada dilakukan penghapusan nilai OEE<0% dan >100% sehingga diperoleh 218 dataset.

### 3.3 Exploratory Data Analysis (EDA)

Tahap selanjutnya adalah melakukan *exploratory data analysis* (EDA). Tujuan dari EDA, yaitu untuk mengetahui karakteristik data dan melihat apakah ada data *outlier* atau tidak dengan visualisasi regplot dan boxplot. Data yang akan digunakan untuk melakukan proses EDA ada sebanyak 218 dataset.

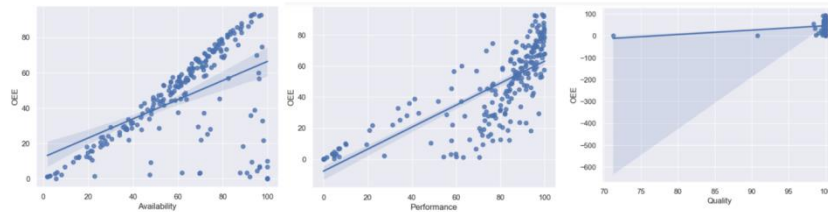


1. Tipe data

Pengecekan tipe data pada variabel *availability*, *performance*, *quality*, dan OEE dengan Program Python diperoleh data yang telah berbentuk numerik sehingga dapat mempermudah peneliti dalam tahap regresi.

2. Visualisasi data

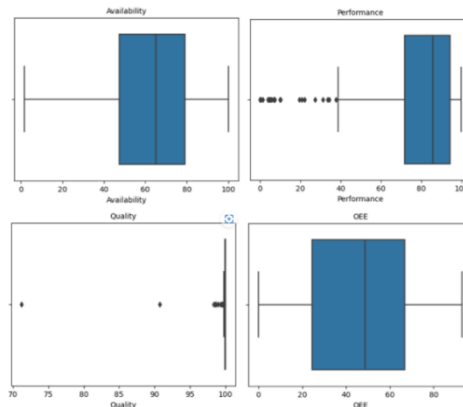
Hasil visualisasi data dengan regplot yang ditampilkan pada Gambar 1.



**Gambar 1. Garfik Antara OEE Dengan Setiap Variabel**

Berdasarkan Gambar 1, variabel yang mempengaruhi nilai OEE yaitu *availability*, dan *performance*. Sedangkan variabel yang tidak mempengaruhi nilai OEE yaitu *quality*. Adapun visualisasi data dengan boxplot untuk setiap variabel ditampilkan pada Gambar 2.

**Gambar 2. Boxplot Setiap Variabel**



Berdasarkan Gambar 2, variabel yang memiliki *outlier* yaitu *performance* dan *quality*. Sehingga perlu menghapus *outlier* agar data bersih dan dapat digunakan dalam penelitian.

**3.4 Penghapusan Outlier**

Penghapusan *outlier* dilakukan agar dataset dapat digunakan dalam proses analisis data. Hasil penghapusan *outlier* yang disajikan pada Tabel 5.

**Tabel 5. Dataset Tanpa Outlier**

No	Observasi Ke-	Availability	Performance	Quality	OEE
1	6	21,11	75,48	99,99	15,93
2	9	62,69	76,25	100,00	47,80
3	12	53,75	98,00	100,00	52,68
4	14	53,48	84,16	99,98	45,00

No	Observasi Ke-	Availability	Performance	Quality	OEE
5	15	60,34	92,35	100,00	55,72
...	...	...	...	...	...
139	669	57,03	77,48	99,97	44,17
140	670	72,62	88,36	100,00	64,16
141	674	79,91	95,85	100,00	76,60
142	678	82,01	97,59	100,00	80,04
143	683	79,25	91,05	100,00	72,15

Berdasarkan Tabel 5, sebanyak 218 dataset yang ada dilakukan proses penghapusan *Outlier* sehingga diperoleh 141 dataset yang tidak memiliki data *outlier*. Jadi, 141 dataset inilah yang akan digunakan untuk memperoleh model regresi.

### 3.5 Pemilihan Variabel

Variabel bebas dan variabel takbebas dapat ditentukan dengan melihat nilai korelasi dari dataset. Jika nilai korelasi mendekati positif satu dan negatif satu maka semakin kuat hubungan antara variabel bebas dan takbebas. Nilai korelasi untuk setiap variabel yang ditampilkan pada Tabel 6 :

**Tabel 6. Nilai Korelasi**

	Availability	Performance	Quality	OEE
Availability	1.00	0.58	0.27	0.97
Performance	0.58	1.00	0.12	0.73
Quality	0.27	0.12	1.00	0.25
OEE	0.97	0.73	0.25	1.00

Berdasarkan Tabel 6, dapat dilihat bahwa nilai korelasi yang paling mempengaruhi OEE adalah *availability* dan *performance* dengan nilai korelasi *availability* yaitu 0,97 dan *performance* yaitu 0,73. Artinya nilai korelasi kedua variabel tersebut mendekati satu, hal ini berarti terdapat keterkaitan yang kuat antara variabel OEE dengan variabel *availability* dan *performance*.

### 3.6 Penerapan Regresi

Perhitungan regresi linear menggunakan 141 dataset dibagi menjadi dua, yaitu data training dan data testing dengan rasio 7:3. Variabel yang digunakan adalah OEE sebagai variabel takbebas sedangkan *availability* dan *performance* sebagai variabel bebas. Proses penerapan regresi linear berganda sebagai berikut :

1. Membagi 141 dataset menjadi dua diperoleh 100 data *training* dan 43 data *testing*.
2. Melatih data *training* yang berjumlah 100 data diperoleh nilai koefisien *intercept* dan *slope* ditampilkan pada Tabel 7.

**Tabel 7. Nilai Koefisien**

Koefisien	Nilai
<i>intercept</i>	-53,72
Variabel <i>availability</i>	0,87
Variabel <i>performance</i>	0,63

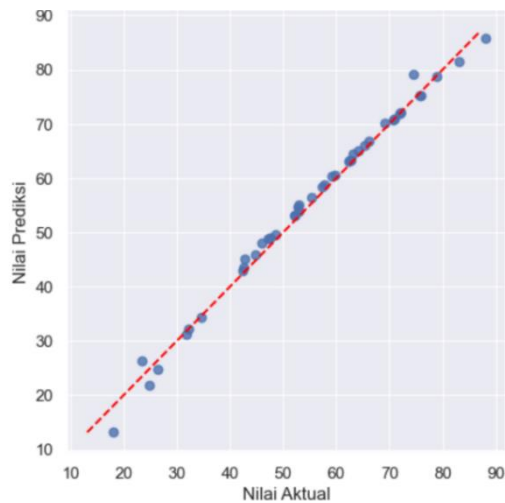
Berdasarkan Tabel 7, diperoleh model regresi linear berganda  $y = -53,72 + 0,87X_1 + 0,63X_2$ .

3. Menguji data *testing* yang berjumlah 43 data sehingga diperoleh nilai prediksi yang disajikan pada Tabel 8.

**Tabel 8. Nilai Prediksi**

No	Nilai Aktual	Nilai Prediksi
1	42,70	45,15
2	47,56	48,99
3	59,65	60,55
4	52,11	53,14
5	75,92	75,23
...	...	...
39	62,38	63,15
40	62,62	63,36
41	42,47	42,94
42	45,96	47,94
43	66,17	66,74

Perbandingan antara nilai aktual dengan nilai prediksi pada Tabel 8 tidak terdapat perbedaan hasil yang jauh dan hasil tersebut disajikan pada Gambar 3.

**Gambar 3. Grafik Antara Nilai Aktual Dan Nilai Prediksi**

Adapun hasil perhitungan *error* ditampilkan pada Tabel 9.

**Tabel 9. Nilai Error**

No	Nilai Aktual	Nilai Prediksi	Error
1	42,70	45,15	-2,45
2	47,56	48,99	-1,43
3	59,65	60,55	-0,89
4	52,11	53,14	-1,03
5	75,92	75,23	0,69
...	...	...	...
39	62,38	63,15	-0,76
40	62,62	63,36	-0,74
41	42,47	42,94	-0,47
42	45,96	47,94	-1,99
43	66,17	66,74	-0,57



Hasil perhitungan eror antara nilai aktual dengan nilai prediksi pada Tabel 9 memperoleh eror yang kecil.

### 3.7 Evaluasi model

Hasil evaluasi model dengan MAPE yaitu 3%. Berdasarkan skala penilaian akurasi MAPE, hasil yang diperoleh berada pada *range* < 10% maka dapat dikatakan bahwa hasil akurasi memiliki model peramalan sangat baik.

## 4. Kesimpulan

Berdasarkan hasil dan pembahasan di atas, dapat disimpulkan bahwa dalam proses penghapusan *outlier* diperoleh 141 dataset yang sesuai dengan format data dan dataset inilah yang digunakan sehingga diperoleh model regresi berganda yaitu  $y = -53,72 + 0,87x_1 + 0,63x_2$ . Dengan menggunakan data *training* dan data *testing* rasio 7:3 diperoleh MAPE prediksi 3% yang memiliki model peramalan sangat baik. Artinya model regresi berganda yang diperoleh dapat digunakan untuk *forecasting* nilai OEE di masa yang akan datang sehingga kinerja mesin dari manufaktur makanan di Indonesia tersebut dapat bekerja secara efektif dan optimal. selain itu, dapat memperoleh target dan hasil yang sesuai dengan yang diinginkan perusahaan.

## REFERENSI

- [1] Mardianto, E. G. (2016). Analisis Faktor-faktor Yang Mempengaruhi Penyerapan Tenaga Kerja Pada Sektor Industri Manufaktur di Kabupaten Bandung Tahun 2001-2014. In *Dipublikasikan*.
- [2] Nakajima, Seiichi. (1988). Introduction to Total Productive Maintenance. *Student Study Guide, TPM100*.
- [3] Badawy, M., El-Aziz, A. A. A., Idress, A. M., Hefny, H., & Hossam, S. (2016). A survey on exploring key performance indicators. *Jurnal Inovasi Penelitian (JIP)*, 1(1–2), 47–52. <https://doi.org/10.1016/j.fcij.2016.04.001>.
- [4] Wiyatno, T. N., M. Fatchan, and A. Firmansyah 2018. “Implementasi Metode Overall Equipment Effectiveness ( Oee ) Guna Mengukur Efektivitas Mesin Produksi,” pp. 559–566.
- [5] Rahmadhani, D. F., Taroepatjeka, H., & Fitria, L. (2014). Usulan Peningkatan Efektivitas Mesin Cetak Manual Menggunakan Metode Overall Equipment Effectiveness (OEE) (Studi Kasus Di Perusahaan Kerupuk TTN). *Jurnal Teknik Industri Ienas*, 02(2338–5081), 156–165.
- [6] Goldberg, D. E., & Holland, J. H. (1988). Genetic Algorithms and Machine Learning. In *Machine Learning* (Vol. 3, Issue 2). <https://doi.org/10.1023/A:1022602019183>
- [7] Daqiqil, I. (2021). *Machine Learning (Teori, Studi Kasus dan Implementasi Menggunakan python* (Edisi 1). UR Press.
- [8] Musyafa'ah, & Sofiana, A. (2022). Analysis of Total Productive Maintenance (TPM) Application Using Overall Equipment Effectiveness (OEE) and Six Big Losses on Disamatic Machine PT. XYZ. *Opsi*, 15(1), 56.
- [9] Ayuni, G. N., & Fitriana, D. (2019). Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti Pada PT XYZ. *Jurnal Telematika*, 14(2), 79–86. <https://journal.ithb.ac.id/telematika/article/view/321>
- [10] Radhi, M., Amalia, Sitompul, D. R. H., Sinurat, S. H., & Indra, E. (2022). Analisis Big Data Dengan Metode Exploratory Data Analysis (EDA) Dan Metode Visualisasi Menggunakan Jupyter Notebook. *Jurnal Sistem Informasi Dan Ilmu Komputer Prima (JUSIKOM PRIMA)*, 4(2), 23–27. <https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v4i2.2475>
- [11] Sholeh, M., Suraya, S., & Andayati, D. (2022). Machine Linear untuk Analisis Regresi Linier Biaya Asuransi Kesehatan dengan Menggunakan Python Jupyter Notebook. *JEPIN (Jurnal Edukasi Dan Penelitian Informatika)*, 8(1), 20–27. <https://jurnal.untan.ac.id/index.php/jepin/article/view/48822>
- [12] Dama, H. R. A., Supianto, A. A., & Setiawan, N. Y. (2021). Analisis Penggunaan Model Regresi untuk Prediksi Penjualan Spare Part pada AHASS Nur Andhita Grogol. 5(12), 5591–5603.
- [13] Boediono, & Koster, W. (2008). *Teori dan Aplikasi Statistika dan Probabilitas* (L. Suryani (ed.); Edisi Keem). PT Remaja Rosdakarya.
- [14] Supranto. (1988). *Statistik Teori dan Aplikasi Jilid 2* (Ju. Mulyadi (ed.); Edisi Keli). Erlangga.
- [15] Lawrence, K. D., Klimberg, R. K., & Lawrence, S. M. (2009). Fundamentals of Forecasting Using Excel. In *Australian Health Service Alliance* (Issue 1).