

## IMPLEMENTASI ALGORITMA *RANDOM FOREST* UNTUK PREDIKSI KETERLAMBATAN PRODUKSI PADA PT PELANGI GRAFIKA

Dennis Ichwan Pratama<sup>1</sup>, Nilma<sup>2</sup>

Program Studi Teknik Informatika, Fakultas Teknik dan Ilmu Komputer  
Universitas Indraprasta PGRI  
Jalan Raya Tengah No 80, Kelurahan Gedong, Pasar Rebo, Jakarta Timur  
pratamaichwan19@gmail.com<sup>1</sup>, nilma23juli@gmail.com<sup>2</sup>

### Abstrak

Keterlambatan produksi merupakan salah satu permasalahan yang dapat memengaruhi efektivitas proses operasional serta kepuasan pelanggan pada industri percetakan. PT Pelangi Grafika masih melakukan identifikasi potensi keterlambatan produksi berdasarkan pengalaman sehingga diperlukan suatu sistem yang mampu memberikan prediksi secara lebih objektif. Penelitian ini bertujuan mengimplementasikan algoritma *Random Forest* pada aplikasi berbasis web untuk memprediksi keterlambatan produksi berdasarkan data historis perusahaan. Dataset yang digunakan terdiri atas atribut jumlah pesanan, jenis produk, *deadline*, estimasi waktu produksi, jumlah mesin, jumlah operator, tingkat kesulitan, jumlah revisi, dan antrian produksi. Tahapan penelitian meliputi *preprocessing* data, pembagian data menjadi *training* dan *testing*, pembentukan model *Random Forest*, serta evaluasi menggunakan *Confusion Matrix* dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa model yang dibangun mampu memberikan performa klasifikasi yang sangat baik dalam memprediksi status keterlambatan produksi. Sistem yang dikembangkan juga mampu membantu perusahaan dalam mengidentifikasi potensi keterlambatan secara lebih cepat sehingga dapat mendukung proses pengambilan keputusan.

**Kata Kunci:** *Random Forest*, Data Mining, Prediksi, Keterlambatan Produksi, *Machine Learning*.

### Abstract

Production delays are one of the problems that can affect operational effectiveness and customer satisfaction in the printing industry. PT Pelangi Grafika still identifies potential production delays based on experience, making it necessary to develop a system capable of providing more objective predictions. This study aims to implement the *Random Forest* algorithm in a web-based application to predict production delays using historical production data. The dataset consists of order quantity, product type, *deadline*, estimated production time, number of machines, number of operators, difficulty level, number of revisions, and production queue. The research stages include data preprocessing, training and testing data splitting, *Random Forest* model construction, and model evaluation using a *Confusion Matrix* with *accuracy*, *precision*, *recall*, and *F1-score* metrics. The results indicate that the proposed model achieves excellent classification performance in predicting production delay status. Furthermore, the developed application assists the company in identifying potential production delays at an earlier stage, thereby supporting more effective decision-making.

**Keyword:** *Random Forest*, Data Mining, Production Delay Prediction, *Machine Learning*.

### PENDAHULUAN

Perkembangan teknologi informasi mendorong perusahaan memanfaatkan *data mining* dan *machine learning* untuk mendukung pengambilan keputusan berdasarkan data historis (Abijono et al., 2021). Salah satu penerapannya adalah prediksi keterlambatan produksi sehingga perusahaan dapat melakukan tindakan antisipatif sebelum keterlambatan terjadi (Nurhalizah, 2024). Metode *data mining* mampu membantu proses prediksi sebagai pendukung pengambilan keputusan (Susanto & Suryadi, 2022). Salah satu algoritma *supervised machine learning* yang banyak digunakan untuk permasalahan klasifikasi adalah *Random Forest* karena mampu menghasilkan performa klasifikasi yang baik pada berbagai jenis data (Indrasetyaning et al., 2021). *Random Forest* membangun sejumlah *decision tree* dan menentukan hasil klasifikasi melalui mekanisme *majority voting*,

sehingga memiliki kemampuan yang baik dalam meningkatkan akurasi serta mengurangi risiko *overfitting* (Kumar & Sinha, 2023).

PT Pelangi Grafika merupakan perusahaan yang bergerak di bidang percetakan dan melayani berbagai kebutuhan cetak, seperti banner, stiker, brosur, buku, dan produk percetakan lainnya. Dalam proses produksinya, perusahaan harus menyelesaikan setiap pesanan sesuai dengan *deadline* yang telah disepakati bersama pelanggan. Namun, proses identifikasi potensi keterlambatan produksi masih dilakukan secara manual berdasarkan pengalaman staf produksi tanpa memanfaatkan data historis yang telah dimiliki perusahaan. Akibatnya, perusahaan mengalami kesulitan dalam mengetahui sejak awal pesanan yang berpotensi mengalami keterlambatan. Kondisi tersebut menyebabkan penyusunan jadwal produksi menjadi kurang optimal, menyulitkan proses pengalokasian sumber daya, serta berpotensi menurunkan kepuasan pelanggan apabila pesanan tidak selesai tepat waktu.

Penelitian ini mengimplementasikan algoritma *Random Forest* pada aplikasi berbasis web untuk memprediksi keterlambatan produksi di PT Pelangi Grafika. Model prediksi dibangun menggunakan atribut jumlah pesanan, jenis produk, *deadline*, estimasi waktu produksi, jumlah mesin, jumlah operator, tingkat kesulitan, jumlah revisi, dan antrian produksi sebagai variabel yang memengaruhi status keterlambatan produksi. Hasil prediksi kemudian ditampilkan melalui sistem sehingga dapat dimanfaatkan oleh pengguna sebagai dasar dalam melakukan pengambilan keputusan terkait proses produksi.

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan penerapan algoritma *Random Forest* yang diintegrasikan ke dalam aplikasi berbasis web untuk memprediksi keterlambatan produksi di PT Pelangi Grafika. Model dievaluasi menggunakan *Confusion Matrix* dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score* sehingga dapat diketahui tingkat kinerja klasifikasinya. Kontribusi penelitian ini adalah menyediakan sistem prediksi yang dapat membantu perusahaan mengidentifikasi potensi keterlambatan produksi sejak dini sebagai dasar dalam mendukung pengambilan keputusan.

## PENELITIAN RELEVAN

Penelitian-penelitian terdahulu menunjukkan bahwa algoritma *Random Forest* memiliki performa yang baik dalam menyelesaikan berbagai permasalahan klasifikasi dan prediksi. Penelitian Liquori et al. (2025) menerapkan algoritma *Random Forest* untuk mengklasifikasikan rentang harga ponsel berdasarkan spesifikasi teknis dan memperoleh hasil klasifikasi dengan tingkat akurasi yang baik. Munggaran et al. (2025) menggunakan algoritma *Random Forest* untuk memprediksi waktu tunggu pelayanan pasien rawat jalan sehingga dapat mendukung peningkatan kualitas pelayanan rumah sakit. Penelitian Muttogar et al. (2022) memanfaatkan *Random Forest* untuk mengklasifikasikan kualitas air sumur dan menunjukkan bahwa algoritma tersebut mampu menghasilkan performa klasifikasi yang baik berdasarkan karakteristik data yang digunakan. Putri & Wijayanto (2022) melakukan analisis komparasi beberapa algoritma klasifikasi pada kasus *website phishing* dan menyimpulkan bahwa *Random Forest* merupakan salah satu algoritma dengan tingkat akurasi yang tinggi dibandingkan algoritma klasifikasi lainnya. Selain itu, penelitian Syaidatussalihah & Abdurahim (2022) membuktikan bahwa *Random Forest* mampu mengklasifikasikan status kemiskinan dengan hasil yang akurat, sedangkan penelitian Suci Amaliah et al. (2022) menerapkan algoritma *Random Forest* untuk klasifikasi varian minuman kopi dan memperoleh hasil klasifikasi dengan tingkat akurasi yang tinggi. Berdasarkan berbagai penelitian tersebut, dapat disimpulkan bahwa algoritma *Random Forest* memiliki kemampuan yang baik dalam menangani permasalahan klasifikasi maupun prediksi pada berbagai bidang.

Perbedaan penelitian ini dengan penelitian-penelitian terdahulu terletak pada objek penelitian, data yang digunakan, serta implementasi sistem yang dikembangkan. Penelitian ini berfokus pada prediksi keterlambatan produksi pada PT Pelangi Grafika menggunakan data historis produksi yang terdiri atas atribut jumlah pesanan, jenis produk, *deadline*, estimasi waktu produksi, jumlah mesin, jumlah operator, tingkat kesulitan, jumlah revisi, dan antrian produksi. Selain membangun model

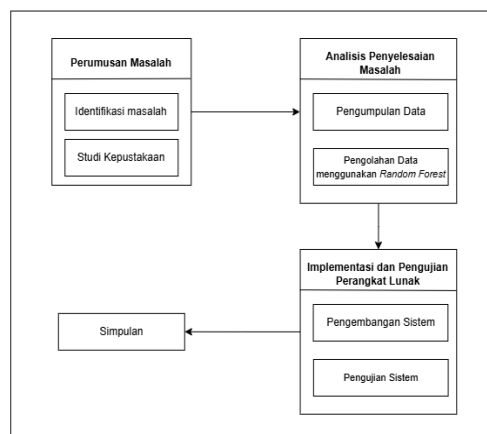
prediksi menggunakan algoritma *Random Forest*, penelitian ini juga mengimplementasikan model tersebut ke dalam aplikasi berbasis web yang dilengkapi dengan fitur pengelolaan data produksi, prediksi keterlambatan, riwayat prediksi, serta penyajian laporan, sehingga dapat membantu perusahaan dalam mengidentifikasi potensi keterlambatan produksi secara lebih cepat dan objektif.

## METODE PENELITIAN

Penelitian ini menerapkan algoritma *Random Forest* untuk memprediksi keterlambatan produksi pada PT Pelangi Grafika. Algoritma *Random Forest* dipilih karena memiliki kemampuan klasifikasi yang baik pada berbagai jenis data (Putri & Wijayanto, 2022). Tahapan penelitian dimulai dengan pengumpulan data historis produksi yang diperoleh melalui observasi dan dokumentasi perusahaan. Dataset yang digunakan terdiri atas atribut jumlah pesanan, jenis produk, *deadline*, estimasi waktu produksi, jumlah mesin, jumlah operator, tingkat kesulitan, jumlah revisi, antrian produksi, serta status keterlambatan sebagai variabel target.

Tahap berikutnya adalah *preprocessing* data yang meliputi pemeriksaan kelengkapan data, penanganan data yang tidak valid, serta transformasi atribut kategorikal menggunakan metode *Label Encoding*. Selanjutnya dataset dibagi menjadi data *training* dan data *testing* menggunakan metode *train-test split* untuk membangun dan menguji model (Suci Amaliah et al., 2022).

Model klasifikasi dibangun menggunakan algoritma *Random Forest*. Algoritma ini membentuk sejumlah *decision tree* berdasarkan data latih dan menentukan hasil prediksi melalui mekanisme *majority voting* (Syaidatussalihah & Abdurahim, 2022). Selama proses pembentukan pohon, pemilihan atribut dilakukan menggunakan nilai *Gini Index* untuk memperoleh percabangan terbaik. Model yang telah dibangun kemudian dievaluasi menggunakan data *testing*. Evaluasi dilakukan dengan *Confusion Matrix* untuk memperoleh nilai *accuracy*, *precision*, *recall*, dan *F1-score*. Selain itu, dilakukan analisis *feature importance* untuk mengetahui atribut yang memberikan pengaruh paling besar terhadap hasil prediksi keterlambatan produksi



Gambar 1. Tahapan Penelitian

Gambar 1 menunjukkan tahapan penelitian yang dilakukan dalam pengembangan sistem prediksi keterlambatan produksi. Penelitian diawali dengan perumusan masalah melalui identifikasi masalah dan studi kepustakaan. Selanjutnya dilakukan analisis penyelesaian masalah yang meliputi pengumpulan data serta penerapan algoritma *Random Forest*. Tahap berikutnya adalah implementasi dan pengujian perangkat lunak, kemudian penelitian diakhiri dengan penarikan simpulan berdasarkan hasil yang diperoleh.

## Rumus Perhitungan:

Accuracy: 
$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Precision:  $Precision = \frac{TP}{TP+FP}$

Recall:  $Recall = \frac{TP}{TP+FN}$

F1-Score:  $\frac{2 \times Precision \times Recall}{Precision + Recall}$

## HASIL DAN PEMBAHASAN

### Dataset

Penelitian ini menggunakan dataset historis produksi PT Pelangi Grafika sebagai data utama dalam proses pembangunan model prediksi. Dataset terdiri atas atribut jumlah pesanan, jenis produk, *deadline*, estimasi waktu produksi, jumlah mesin, jumlah operator, tingkat kesulitan, jumlah revisi, antrian produksi, serta status keterlambatan sebagai variabel target. Data tersebut selanjutnya digunakan dalam proses pelatihan dan pengujian model *Random Forest*.

Tabel 1. Dataset

| No | Atribut                 | Keterangan                            |
|----|-------------------------|---------------------------------------|
| 1  | Jumlah Pesanan          | Jumlah pesanan yang diproduksi        |
| 2  | Jenis Produk            | Jenis produk percetakan               |
| 3  | Deadline                | Batas waktu penyelesaian              |
| 4  | Estimasi Waktu Produksi | Perkiraan waktu produksi              |
| 5  | Jumlah Mesin            | Jumlah mesin yang digunakan           |
| 6  | Jumlah Operator         | Jumlah operator produksi              |
| 7  | Tingkat Kesulitan       | Tingkat kesulitan pekerjaan           |
| 8  | Jumlah Revisi           | Jumlah revisi pesanan                 |
| 9  | Antrian Produksi        | Jumlah pekerjaan dalam antrian        |
| 10 | Status Keterlambatan    | Label klasifikasi (Telat/Tidak Telat) |

### Preprocessing Data

#### Pemeriksaan Struktur Dataset

```
df.info()
[4] ✓ 0.0s
...
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 600 entries, 0 to 599
Data columns (total 11 columns):
#   Column              Non-Null Count  Dtype
---  -
0   tanggal_pesanan     600 non-null   object
1   jumlah_pesanan      600 non-null   int64
2   jenis_produk        600 non-null   object
3   deadline_hari       600 non-null   int64
4   estimasi_waktu_produksi 600 non-null   int64
5   jumlah_mesin        600 non-null   int64
6   jumlah_operator     600 non-null   int64
7   tingkat_kesulitan   600 non-null   object
8   jumlah_revisi       600 non-null   int64
9   antrian_produksi    600 non-null   int64
10  status_keterlambatan 600 non-null   object
dtypes: int64(7), object(4)
memory usage: 51.7+ KB
```

Gambar 2. Struktur Dataset

*Output* `df.info()` menunjukkan informasi mengenai struktur dataset, seperti jumlah data, nama atribut, tipe data setiap atribut, serta jumlah nilai yang tidak kosong (*non-null*). Informasi ini

digunakan untuk memastikan bahwa seluruh atribut telah terbaca dengan benar dan siap digunakan pada proses selanjutnya.

### Menampilkan Data Awal

|   | tanggal_pesanan | jumlah_pesanan | jenis_produk | deadline_hari | estimasi_waktu_produk | jumlah_mesin | jumlah_operator | tingkat_kesulitan | jumlah_revisi | antrian_produk | status_keterlambatan |
|---|-----------------|----------------|--------------|---------------|-----------------------|--------------|-----------------|-------------------|---------------|----------------|----------------------|
| 0 | 01-01-2025      | 39             | Banner       | 4             | 1                     | 1            | 3               | Rendah            | 2             | 19             | Telat                |
| 1 | 01-01-2025      | 29             | Banner       | 1             | 9                     | 3            | 1               | Rendah            | 2             | 10             | Telat                |
| 2 | 02-01-2025      | 15             | Stiker       | 5             | 8                     | 2            | 5               | Rendah            | 0             | 10             | Telat                |
| 3 | 02-01-2025      | 43             | Brosur       | 9             | 2                     | 4            | 1               | Rendah            | 2             | 10             | Tidak Telat          |
| 4 | 03-01-2025      | 8              | Banner       | 1             | 11                    | 2            | 3               | Rendah            | 1             | 1              | Telat                |
| 5 | 04-01-2025      | 21             | Stiker       | 9             | 8                     | 2            | 5               | Sedang            | 0             | 16             | Telat                |
| 6 | 04-01-2025      | 39             | Brosur       | 8             | 9                     | 2            | 3               | Rendah            | 1             | 8              | Telat                |
| 7 | 05-01-2025      | 19             | Buku         | 6             | 1                     | 2            | 5               | Sedang            | 4             | 1              | Tidak Telat          |
| 8 | 05-01-2025      | 23             | Stiker       | 7             | 5                     | 3            | 3               | Rendah            | 3             | 14             | Telat                |
| 9 | 06-01-2025      | 11             | Brosur       | 3             | 8                     | 4            | 2               | Tinggi            | 4             | 15             | Telat                |

Gambar 3. Dataset Awal

### Pemeriksaan Missing Value

|                       | df.isnull().sum() |
|-----------------------|-------------------|
| tanggal_pesanan       | 0                 |
| jumlah_pesanan        | 0                 |
| jenis_produk          | 0                 |
| deadline_hari         | 0                 |
| estimasi_waktu_produk | 0                 |
| jumlah_mesin          | 0                 |
| jumlah_operator       | 0                 |
| tingkat_kesulitan     | 0                 |
| jumlah_revisi         | 0                 |
| antrian_produk        | 0                 |
| status_keterlambatan  | 0                 |
| dtype: int64          |                   |

Gambar 4. Cek Missing Value

Hasil pemeriksaan menunjukkan jumlah nilai kosong pada masing-masing atribut. Seluruh atribut bernilai 0, maka dapat disimpulkan bahwa dataset tidak memiliki *missing value* sehingga tidak diperlukan proses penanganan data kosong dan data dapat langsung digunakan pada tahap berikutnya.

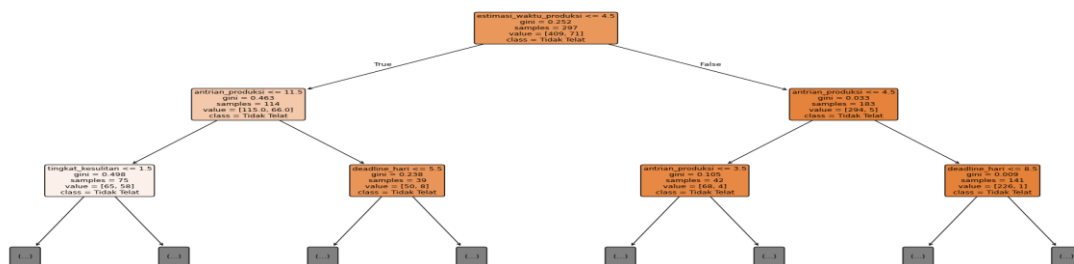
### Label encoding

Tabel 2. Label Encoding

| Atribut              | Kategori    | Nilai Encoding |
|----------------------|-------------|----------------|
| Jenis Produk         | Banner      | 0              |
|                      | Brosur      | 1              |
|                      | Buku        | 2              |
|                      | Stiker      | 3              |
| Tingkat Kesulitan    | Rendah      | 0              |
|                      | Sedang      | 1              |
|                      | Tinggi      | 2              |
| Status Keterlambatan | Tidak Telat | 0              |
|                      | Telat       | 1              |

### Implementasi Algoritma *Random Forest*

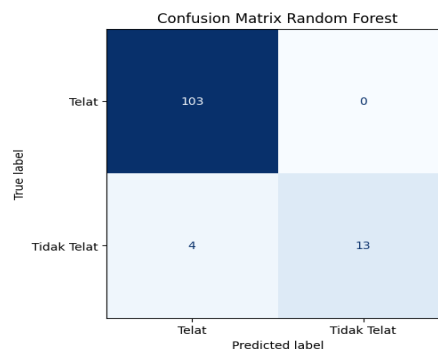
Model prediksi dibangun menggunakan algoritma *Random Forest* berdasarkan data *training*. Algoritma membentuk sejumlah *decision tree* melalui proses *bootstrap sampling* dan *random feature selection*. Pembentukan banyak pohon keputusan melalui proses *bootstrap sampling* dan pemilihan atribut secara acak menjadikan *Random Forest* lebih stabil dalam menangani variasi data (Probst et al., 2019). Selanjutnya, setiap pohon menghasilkan prediksi yang kemudian digabungkan menggunakan mekanisme *majority voting* untuk menentukan hasil akhir berupa status Telat atau Tidak Telat. Untuk memberikan gambaran mengenai proses klasifikasi yang dilakukan model, salah satu *decision tree* tapi tidak keseluruhan karena terlalu panjang ada 100 pohon. Jadi ini hasil pembentukan *Random Forest* ditampilkan pada Gambar dibawah ini.



Gambar 5. *Decision Tree*

### Evaluasi Model

Evaluasi model menggunakan *confusion matrix* dilakukan untuk mengetahui kemampuan model dalam mengklasifikasikan data melalui perbandingan antara hasil prediksi dan data aktual sehingga diperoleh nilai *accuracy*, *precision*, *recall*, dan *F1-score* (Nugroho, 2022). Setelah model berhasil dibangun, dilakukan proses evaluasi menggunakan data *testing*. Evaluasi dilakukan dengan *Confusion Matrix* untuk mengetahui kemampuan model dalam mengklasifikasikan status keterlambatan produksi berdasarkan data yang belum pernah dipelajari sebelumnya.



Gambar 6. *Confusion Matrix*

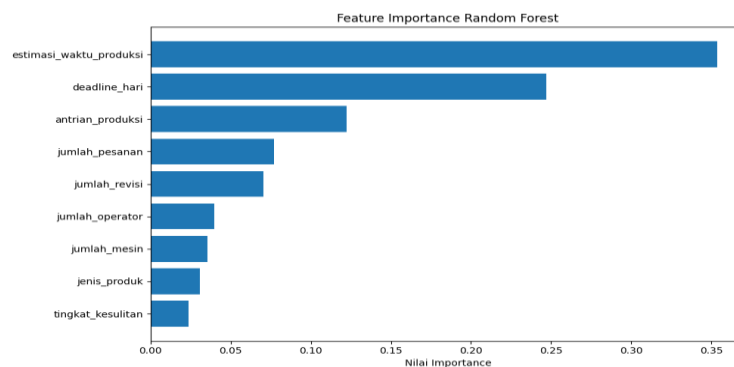
Tabel 3. Hasil Nilai *Confusion Matrix*

| Keterangan          | Nilai |
|---------------------|-------|
| True Positive (TP)  | 103   |
| False Negative (FN) | 0     |
| False Positive (FP) | 4     |
| True Negative (TN)  | 13    |

$$\begin{aligned}
 \text{Accuracy:} &= \frac{TP+TN}{TP+TN+FP+FN} = \frac{103+13}{103+13+4+0} = \frac{116}{120} = 0,9667 \times 100\% = 96,67\% \\
 \text{Precision:} &= \frac{TP}{TP+FP} = \frac{103}{103+4} = \frac{103}{107} = 0,9626 \times 100\% = 96,26\% \\
 \text{Recall:} &= \frac{TP}{TP+FN} = \frac{103}{103+0} = \frac{103}{103} = 1,00 \times 100\% = 100\% \\
 \text{F1-Score:} &= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0,9626 \times 1,00}{0,9626 + 1,00} = \frac{1,9252}{1,9626} = 0,9809 \times 100\% = 98,09\%
 \end{aligned}$$

### Analisis Feature Importance

Selain mengevaluasi performa model, penelitian ini juga menganalisis tingkat pengaruh masing-masing atribut menggunakan *feature importance*. Analisis ini bertujuan untuk mengetahui atribut yang paling berkontribusi terhadap proses prediksi keterlambatan produksi.

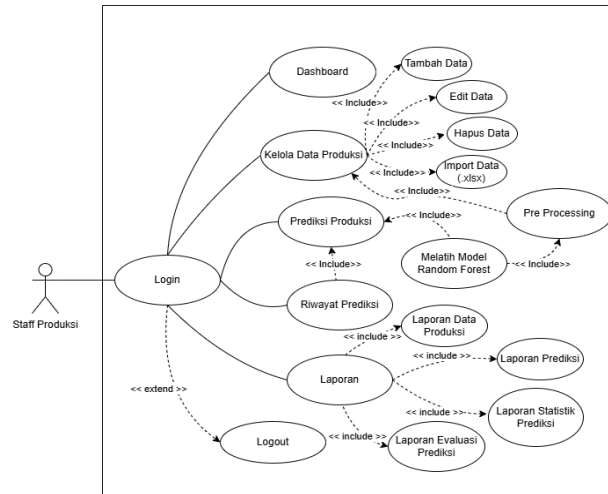


Gambar 7. Feature Importance

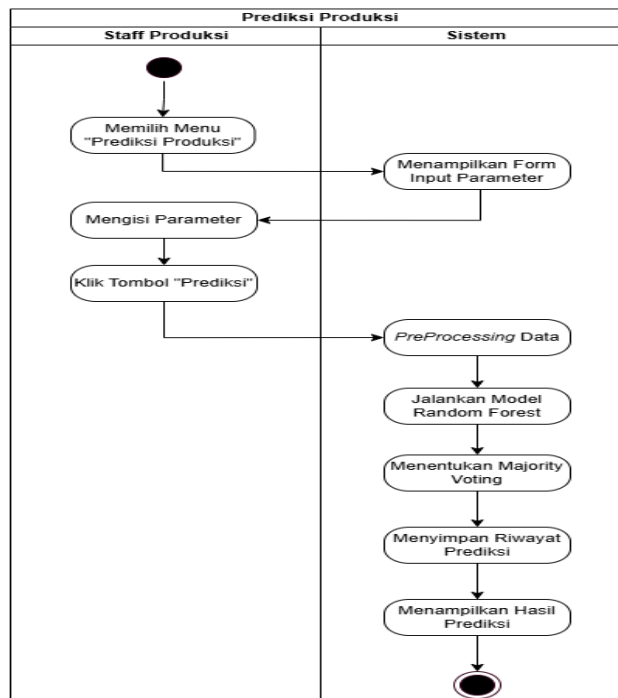
Berdasarkan gambar diatas, diketahui bahwa variabel estimasi waktu produksi memiliki nilai *Feature Importance* tertinggi, sehingga menjadi variabel yang paling berpengaruh dalam proses prediksi keterlambatan produksi. Hal ini menunjukkan bahwa lamanya waktu yang dibutuhkan untuk menyelesaikan suatu pesanan sangat memengaruhi kemungkinan terjadinya keterlambatan.

### Pemodelan Perangkat Lunak

Pemodelan sistem menggunakan *Unified Modeling Language* untuk menggambarkan fungsi dan alur utama aplikasi. *Use Case Diagram* memperlihatkan interaksi Administrator dengan pengelolaan kriteria, subkriteria, kandidat, perbandingan kriteria, penilaian, hasil rekomendasi, dan laporan. *Activity Diagram* menggambarkan alur ketika Administrator melihat hasil pemeringkatan dan memilih laporan yang akan diunduh. Penggunaan UML membantu dokumentasi rancangan dan komunikasi pengembangan sistem (Wayahdi & Ruziq, 2023).

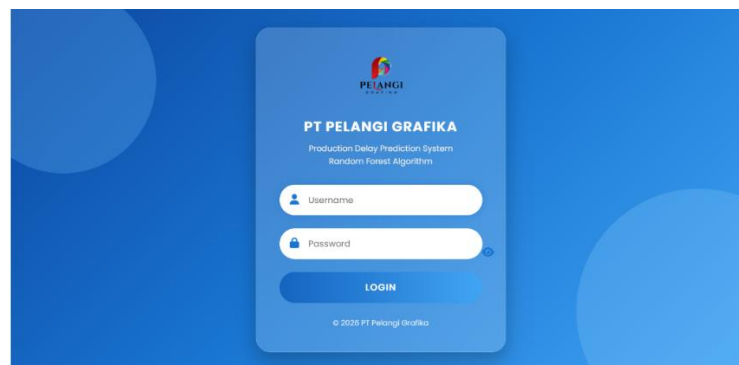


Gambar 8. Use Case Diagram



Gambar 9. Activity Diagram Melakukan Prediksi Produksi

## Tampilan Layar *Login* dan *Dashboard*



Gambar 10. Tampilan Layar *Login*

Tampilan layar *login* merupakan implementasi dari rancangan halaman login yang digunakan sebagai proses autentikasi pengguna. Pengguna harus memasukkan username dan password yang sesuai agar dapat mengakses seluruh fitur yang tersedia pada sistem.



Gambar 11. Tampilan Layar *Dashboard*

Tampilan layar *dashboard* menampilkan informasi utama mengenai sistem setelah pengguna berhasil login. Halaman ini berfungsi sebagai pusat navigasi yang memberikan akses cepat menuju menu data produksi, prediksi, riwayat prediksi, dan laporan.

### Tampilan Layar Riwayat Prediksi Produksi dan Laporan Prediksi Produksi

| No | Tanggal          | Produk | Pesanan | Deadline | Hasil Prediksi | Pengguna | Aksi           |
|----|------------------|--------|---------|----------|----------------|----------|----------------|
| 1  | 02/07/2026 11:58 | Brosur | 78      | 2 Hari   | TELAT          | admin    | [Edit] [Hapus] |
| 2  | 02/07/2026 11:57 | Banner | 120     | 3 Hari   | TELAT          | admin    | [Edit] [Hapus] |
| 3  | 30/06/2026 22:32 | Banner | 50      | 4 Hari   | TELAT          | admin    | [Edit] [Hapus] |
| 4  | 30/06/2026 22:21 | Banner | 50      | 4 Hari   | TELAT          | admin    | [Edit] [Hapus] |
| 5  | 30/06/2026 22:04 | Banner | 50      | 4 Hari   | TELAT          | admin    | [Edit] [Hapus] |

Gambar 12. Tampilan Layar Riwayat Prediksi

Tampilan layar riwayat prediksi menampilkan seluruh hasil prediksi yang telah dilakukan pengguna. Informasi yang tersimpan dapat digunakan sebagai dokumentasi maupun sebagai bahan evaluasi terhadap hasil prediksi yang telah dilakukan sebelumnya.

| No | Tanggal Prediksi | Jenis Produk | Jumlah Pesanan | Deadline | Hasil Prediksi | Progress |
|----|------------------|--------------|----------------|----------|----------------|----------|
| 1  | 21/07/2026 09:51 | Brosur       | 90             | 2 Hari   | Telat          | admin    |
| 2  | 14/07/2026 12:28 | Brosur       | 35             | 4 Hari   | Telat          | admin    |
| 3  | 30/06/2026 12:22 | Brosur       | 50             | 4 Hari   | Telat          | admin    |
| 4  | 30/06/2026 12:11 | Brosur       | 50             | 4 Hari   | Telat          | admin    |
| 5  | 30/06/2026 12:04 | Brosur       | 50             | 4 Hari   | Telat          | admin    |
| 6  | 30/06/2026 11:57 | Brosur       | 45             | 4 Hari   | Telat          | admin    |
| 7  | 30/06/2026 11:49 | Brosur       | 50             | 4 Hari   | Telat          | admin    |

Jakarta, 22 Juli 2026 10:18 WIB  
Staff Produksi,  
YOHANA

Gambar 13. Tampilan Output Laporan Prediksi Produksi

## SIMPULAN

Berdasarkan hasil penelitian, algoritma *Random Forest* berhasil diimplementasikan untuk memprediksi keterlambatan produksi pada PT Pelangi Grafika berdasarkan data historis produksi. Model yang dibangun mampu mengklasifikasikan status produksi ke dalam kategori Telat dan Tidak Telat dengan performa yang sangat baik. Hasil evaluasi menunjukkan nilai *Accuracy* sebesar 96,67%, *Precision* sebesar 96,26%, *Recall* sebesar 100%, dan *F1-Score* sebesar 98,09%, sehingga model dinilai mampu memberikan hasil prediksi yang akurat.

Selain itu, hasil analisis *feature importance* menunjukkan bahwa atribut estimasi waktu produksi, deadline hari, dan antrian produksi merupakan faktor yang paling berpengaruh dalam proses prediksi keterlambatan produksi. Dengan demikian, penerapan algoritma *Random Forest* dapat dimanfaatkan sebagai pendukung pengambilan keputusan dalam mengidentifikasi potensi keterlambatan produksi sejak dini sehingga perusahaan dapat melakukan langkah antisipatif untuk meminimalkan risiko keterlambatan.

## DAFTAR PUSTAKA

- Abijono, H., Santoso, P., & Anggreini, N. L. (2021). Supervised Learning And Unsupervised Learning Algorithm In Data Processing. *G-Tech: Jurnal Teknologi Terapan*, 4(2), 315–318.
- Indrasetianingsih, A., Fitriani, F., & Kusuma, P. J. (2021). *Klasifikasi Indeks Pembangunan Gender di Indonesia Tahun 2020 Menggunakan Supervised Machine Learning Algorithms*.
- Kumar, V., & Sinha, D. (2023). Random Forest Based Classification Techniques: A Review of Recent Applications and Performance Improvements. *International Journal of Advanced Computer Science and Applications*, 14(5), 421–429.
- Nugroho, A. (2022). *Analisa Splitting Criteria Pada Decision Tree dan Random Forest untuk Klasifikasi Evaluasi Kendaraan*. 1(1), 41–49.
- Nurhalizah, R. (2024). *Analisis Supervised dan Unsupervised Learning pada Machine Learning: Systematic Literature Review*.
- Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and Tuning Strategies for Random Forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(3). <https://doi.org/10.1002/widm.1301>
- Putri, N. B., & Wijayanto, A. W. (2022). Analisis Komparasi Algoritma Klasifikasi Data Mining Dalam Klasifikasi Website Phishing Comparative Analysis Of Data Mining Classification Algorithm In Phishing Website Classification. *Komputika*, 11(28), 59–66. <https://doi.org/10.34010/komputika.v11i1.4350>
- Suci Amaliah, Nusrang, M., & Aswi, A. (2022). Penerapan Metode Random Forest Untuk Klasifikasi Varian Minuman Kopi di Kedai Kopi Konijiwa Bantaeng. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 4(3), 121–127. <https://doi.org/10.35580/variansiunm31>
- Susanto, A., & Suryadi, D. (2022). Penerapan Metode Data Mining untuk Prediksi dan Pengambilan Keputusan pada Sistem Informasi. *Jurnal Teknologi Informasi Dan Komunikasi*, 13(2), 45–53.
- Syaidatussalihah, & Abdurahim. (2022). Classification of Poverty Status using the Random Forest Algorithm. *EIGEN MATHEMATICS JOURNAL*, 5(1 SE-Articles), 37–44. <https://doi.org/10.29303/emj.v5i1.133>
- Wayahdi, M. R., & Ruziq, F. (2023). *Pemodelan Sistem Penerimaan Anggota Baru dengan Unified Modeling Language (UML) (Studi Kasus: Programmer Association of Battuta)*. 12, 1514–1521.